



85	15.2	76.0	598	13	US-10-027-632-194904	Sequence 194904,	C 158	14.8	74.0	564	17	US-10-027-632-267664	Sequence 267664,
86	15.2	76.0	598	13	US-10-027-632-194905	Sequence 194905,	C 159	14.8	74.0	564	17	US-10-027-632-267666	Sequence 267666,
87	15.2	76.0	598	13	US-10-027-632-194906	Sequence 194906,	C 160	14.8	74.0	575	13	US-10-027-632-269148	Sequence 269148,
88	15.2	76.0	598	17	US-10-027-632-194905	Sequence 194905,	C 161	14.8	74.0	575	13	US-10-027-632-269149	Sequence 269149,
89	15.2	76.0	603	10	US-09-814-353-16350	Sequence 16350, A	C 162	14.8	74.0	575	13	US-10-027-632-269150	Sequence 269150,
90	15.2	76.0	641	13	US-10-027-632-133093	Sequence 133093,	C 163	14.8	74.0	575	17	US-10-027-632-269148	Sequence 269148,
91	15.2	76.0	641	17	US-10-027-632-133093	Sequence 133093,	C 164	14.8	74.0	575	17	US-10-027-632-269149	Sequence 269149,
92	15.2	76.0	685	17	US-10-027-632-142147	Sequence 142147,	C 165	14.8	74.0	575	17	US-10-027-632-269150	Sequence 269150,
93	15.2	76.0	685	17	US-10-027-632-142147	Sequence 142147,	C 166	14.8	74.0	576	13	US-10-027-632-41732	Sequence 41732, A
94	15.2	76.0	993	17	US-10-424-599-23225	Sequence 23225, A	C 167	14.8	74.0	576	17	US-10-027-632-41732	Sequence 41732, A
95	15.2	76.0	1005	17	US-10-295-027-701	Sequence 701, App	C 168	14.8	74.0	585	13	US-10-027-632-63052	Sequence 63052, A
96	15.2	76.0	1070	13	US-10-098-841-168	Sequence 168, App	C 169	14.8	74.0	585	13	US-10-027-632-63938	Sequence 63938, A
97	15.2	76.0	2258	17	US-10-027-632-103137	Sequence 103137,	C 170	14.8	74.0	585	13	US-10-027-632-297454	Sequence 297454,
98	15.2	76.0	2258	17	US-10-027-632-103137	Sequence 103137,	C 171	14.8	74.0	585	17	US-10-027-632-63052	Sequence 63052, A
99	15.2	76.0	2264	17	US-10-421-554-71	Sequence 71, Appl	C 172	14.8	74.0	585	17	US-10-027-632-63938	Sequence 63938, A
100	15.2	76.0	3828	15	US-10-128-714-2125	Sequence 2125, App	C 173	14.8	74.0	585	17	US-10-027-632-297454	Sequence 297454,
101	15.2	76.0	3853	17	US-10-264-049-681	Sequence 681, App	C 174	14.8	74.0	620	16	US-10-417-510-21	Sequence 21, Appl
102	15.2	76.0	4079	15	US-10-128-714-1125	Sequence 1125, App	C 175	14.8	74.0	620	16	US-10-723-860-1656	Sequence 1656, App
103	15.2	76.0	4308	15	US-10-128-714-1125	Sequence 1125, App	C 176	14.8	74.0	631	13	US-10-027-632-79613	Sequence 79613, A
104	15.2	76.0	4627	15	US-10-128-714-6125	Sequence 6125, App	C 177	14.8	74.0	631	13	US-10-027-632-301275	Sequence 301275, A
105	15.2	76.0	6079	15	US-10-128-714-125	Sequence 125, App	C 178	14.8	74.0	631	17	US-10-027-632-79613	Sequence 79613, A
106	15.2	76.0	6089	18	US-10-758-672A-5	Sequence 5, Appl	C 179	14.8	74.0	631	17	US-10-027-632-301275	Sequence 301275, A
107	15.2	76.0	6627	15	US-10-128-714-5125	Sequence 5125, App	C 180	14.8	74.0	635	17	US-10-424-599-75120	Sequence 75120, A
108	15.2	76.0	30072	13	US-10-087-192-457	Sequence 457, App	C 181	14.8	74.0	654	17	US-10-424-599-14315	Sequence 14315, A
109	15.2	76.0	73725	17	US-10-052-482-196	Sequence 196, App	C 182	14.8	74.0	654	18	US-10-767-701-25770	Sequence 25770, A
110	15.2	76.0	80423	18	US-10-367-094-41	Sequence 41, Appl	C 183	14.8	74.0	705	13	US-10-027-632-267668	Sequence 267668,
111	15.2	76.0	120144	18	US-10-322-281-143	Sequence 143, App	C 184	14.8	74.0	705	16	US-10-417-510-3	Sequence 3, Appl
112	15.2	76.0	143068	9	US-09-967-768A-316	Sequence 316, App	C 185	14.8	74.0	705	17	US-10-027-632-267668	Sequence 267668,
113	15.2	76.0	309996	18	US-10-719-993-6785	Sequence 6785, App	C 186	14.8	74.0	728	13	US-10-027-632-19728	Sequence 19728, A
114	15.2	76.0	684973	9	US-09-263-959-1	Sequence 1, Appl	C 187	14.8	74.0	728	17	US-10-027-632-19728	Sequence 19728, A
115	15	75.0	624	18	US-10-437-963-63738	Sequence 63738, A	C 188	14.8	74.0	741	18	US-10-425-115-150893	Sequence 150893, A
116	15	75.0	843	18	US-10-437-963-69461	Sequence 69461, A	C 189	14.8	74.0	741	18	US-10-425-115-150893	Sequence 150893, A
117	15	75.0	1048	17	US-10-424-599-118822	Sequence 118822, A	C 190	14.8	74.0	932	17	US-10-425-114-4836	Sequence 4836, App
118	14.8	74.0	25	19	US-10-719-900-51824	Sequence 51824, A	C 191	14.8	74.0	946	17	US-10-425-114-4836	Sequence 4836, App
119	14.8	74.0	25	19	US-10-719-900-192211	Sequence 192211,	C 192	14.8	74.0	955	18	US-10-684-422-17	Sequence 17, Appl
120	14.8	74.0	25	19	US-10-719-900-908657	Sequence 908657,	C 193	14.8	74.0	957	18	US-10-774-355A-359	Sequence 359, App
121	14.8	74.0	201	18	US-10-719-993-12788	Sequence 12788, A	C 194	14.8	74.0	960	19	US-10-425-115-110451	Sequence 110451,
122	14.8	74.0	339	18	US-10-767-701-16424	Sequence 16424, A	C 195	14.8	74.0	960	19	US-10-774-355A-408	Sequence 408, App
123	14.8	74.0	351	17	US-10-424-599-52826	Sequence 52826, A	C 196	14.8	74.0	991	18	US-10-425-115-35151	Sequence 35151, A
124	14.8	74.0	413	18	US-10-425-115-162973	Sequence 162973,	C 197	14.8	74.0	993	18	US-10-437-963-43376	Sequence 43376, A
125	14.8	74.0	464	13	US-10-027-632-311653	Sequence 311653,	C 198	14.8	74.0	1014	18	US-10-425-115-33505	Sequence 33505, A
126	14.8	74.0	464	13	US-10-027-632-311653	Sequence 311653,	C 199	14.8	74.0	1042	17	US-10-424-599-98223	Sequence 98223, A
127	14.8	74.0	464	17	US-10-027-632-311654	Sequence 311653,	C 200	14.8	74.0	1087	18	US-10-425-115-35152	Sequence 35152, A
128	14.8	74.0	464	17	US-10-027-632-311654	Sequence 311654,	C 201	14.8	74.0	1172	18	US-10-739-930-5110	Sequence 5110, App
129	14.8	74.0	465	17	US-10-424-599-137924	Sequence 137924, A	C 202	14.8	74.0	1311	17	US-10-369-493-47015	Sequence 47015, A
130	14.8	74.0	497	17	US-10-242-535A-24422	Sequence 24422, A	C 203	14.8	74.0	1565	17	US-10-425-114-27541	Sequence 27541, A
131	14.8	74.0	497	17	US-10-085-783A-24422	Sequence 24422, A	C 204	14.8	74.0	1815	18	US-10-425-115-175250	Sequence 175250, A
132	14.8	74.0	497	18	US-10-425-115-40003	Sequence 40003, A	C 205	14.8	74.0	1873	9	US-09-834-975-756	Sequence 756, App
133	14.8	74.0	512	13	US-10-027-632-267660	Sequence 267660,	C 206	14.8	74.0	1873	9	US-09-834-975-757	Sequence 757, App
134	14.8	74.0	512	13	US-10-027-632-267665	Sequence 267665,	C 207	14.8	74.0	2191	10	US-09-931-733-1	Sequence 1, Appl
135	14.8	74.0	512	13	US-10-027-632-267667	Sequence 267667,	C 208	14.8	74.0	2191	10	US-10-437-963-33304	Sequence 33304, A
136	14.8	74.0	512	17	US-10-027-632-267660	Sequence 267660,	C 209	14.8	74.0	3512	13	US-10-027-632-114278	Sequence 114278,
137	14.8	74.0	512	17	US-10-027-632-267665	Sequence 267665,	C 210	14.8	74.0	3512	13	US-10-027-632-114279	Sequence 114279,
138	14.8	74.0	512	17	US-10-027-632-267667	Sequence 267667,	C 211	14.8	74.0	3512	17	US-10-027-632-114278	Sequence 114278,
139	14.8	74.0	519	9	US-09-918-995-5944	Sequence 5944, App	C 212	14.8	74.0	3512	9	US-10-027-632-114279	Sequence 114279,
140	14.8	74.0	529	9	US-09-864-761-7787	Sequence 7787, App	C 213	14.8	74.0	4836	9	US-09-764-869-1735	Sequence 1735, App
141	14.8	74.0	538	18	US-10-767-701-18075	Sequence 18075, A	C 214	14.8	74.0	7537	17	US-10-091-504-1735	Sequence 1735, App
142	14.8	74.0	554	13	US-10-027-632-223235	Sequence 223235,	C 215	14.8	74.0	7537	17	US-10-027-577-1735	Sequence 1735, App
143	14.8	74.0	554	13	US-10-027-632-223236	Sequence 223236,	C 216	14.8	74.0	40304	11	US-09-997-722-280	Sequence 280, App
144	14.8	74.0	554	13	US-10-027-632-223237	Sequence 223237,	C 217	14.8	74.0	41322	15	US-10-024-396-13	Sequence 13, Appl
145	14.8	74.0	554	13	US-10-027-632-223238	Sequence 223238,	C 218	14.8	74.0	44990	17	US-10-052-482-217	Sequence 217, App
146	14.8	74.0	554	17	US-10-027-632-223235	Sequence 223235,	C 219	14.8	74.0	4836	9	US-09-927-602-38	Sequence 38, Appl
147	14.8	74.0	554	17	US-10-027-632-223236	Sequence 223236,	C 220	14.8	74.0	86574	18	US-10-775-169-173	Sequence 173, App
148	14.8	74.0	554	17	US-10-027-632-223237	Sequence 223237,	C 221	14.8	74.0	87869	18	US-10-717-597-194	Sequence 194, App
149	14.8	74.0	554	17	US-10-027-632-223238	Sequence 223238,	C 222	14.8	74.0	106938	18	US-10-322-281-566	Sequence 566, App
150	14.8	74.0	560	13	US-10-027-632-187475	Sequence 187475,	C 223	14.8	74.0	160552	18	US-10-697-828-11	Sequence 11, Appl
151	14.8	74.0	560	17	US-10-027-632-187476	Sequence 187476,	C 224	14.8	74.0	175338	13	US-10-087-192-1861	Sequence 1861, App
152	14.8	74.0	560	17	US-10-027-632-187475	Sequence 187475,	C 225	14.8	74.0	193303	15	US-10-081-327-37	Sequence 37, Appl
153	14.8	74.0	560	17	US-10-027-632-187476	Sequence 187476,	C 226	14.8	74.0	193303	15	US-10-081-327-37	Sequence 37, Appl
154	14.8	74.0	564	13	US-10-027-632-267661	Sequence 267661,	C 227	14.8	74.0	197997	9	US-09-822-246-3	Sequence 44, Appl
155	14.8	74.0	564	13	US-10-027-632-267664	Sequence 267664,	C 228	14.8	74.0	197997	18	US-10-469-028-3	Sequence 3, Appl
156	14.8	74.0	564	13	US-10-027-632-267666	Sequence 267666,	C 229	14.8	74.0	219352	18	US-10-322-281-45	Sequence 45, Appl
157	14.8	74.0	564	17	US-10-027-632-267661	Sequence 267661,	C 230	14.8	74.0	277616	18	US-10-367-094-83	Sequence 83, Appl

231	14.8	74.0	421609	18	US-10-367-094-122	Sequence 122, App	304	14.4	72.0	5253	14	US-10-080-943-3	Sequence 3, Appli
232	14.8	74.0	775062	18	US-10-719-993-6844	Sequence 6844, App	305	14.4	72.0	5549	18	US-10-723-860-5037	Sequence 5037, Ap
233	14.8	74.0	1223197	13	US-10-027-632-179264	Sequence 179264, A	306	14.4	72.0	8764	10	US-09-764-891-10019	Sequence 10019, A
234	14.8	74.0	1223197	17	US-10-027-632-179264	Sequence 179264, A	307	14.4	72.0	10445	10	US-09-764-891-6380	Sequence 6380, Ap
235	14.8	74.0	1691139	14	US-10-067-514-1	Sequence 1, Appli	308	14.4	72.0	14747	16	US-10-286-926-42	Sequence 42, Appl
236	14.8	74.0	1691139	17	US-10-419-723-1	Sequence 1, Appli	309	14.4	72.0	18340	16	US-10-240-376A-130	Sequence 130, App
237	14.4	72.0	129	17	US-10-424-599-69773	Sequence 69773, A	310	14.4	72.0	21704	15	US-10-017-161-709	Sequence 709, App
238	14.4	72.0	210	18	US-10-674-124A-12661	Sequence 12661, A	311	14.4	72.0	21704	15	US-10-292-798-621	Sequence 621, App
239	14.4	72.0	232	17	US-10-424-599-18683	Sequence 18683, A	312	14.4	72.0	25346	18	US-10-601-700-15	Sequence 15, Appl
240	14.4	72.0	271	16	US-10-029-386-14471	Sequence 14471, A	313	14.4	72.0	32134	17	US-10-242-355-936	Sequence 936, App
241	14.4	72.0	296	18	US-10-723-860-3871	Sequence 3871, App	314	14.4	72.0	32192	17	US-10-242-355-937	Sequence 937, App
242	14.4	72.0	302	17	US-10-424-599-104449	Sequence 104449, A	315	14.4	72.0	35460	18	US-10-322-696-94	Sequence 94, Appl
243	14.4	72.0	315	18	US-10-357-930-6846	Sequence 6846, App	316	14.4	72.0	96276	18	US-10-723-860-451	Sequence 451, App
244	14.4	72.0	328	18	US-10-357-930-6547	Sequence 6547, App	317	14.4	72.0	96596	17	US-10-052-482-10	Sequence 10, Appl
245	14.4	72.0	364	18	US-10-674-124A-21095	Sequence 21095, A	318	14.4	72.0	168198	18	US-10-322-696-88	Sequence 88, Appl
246	14.4	72.0	366	17	US-10-045-735A-23201	Sequence 23201, A	319	14.4	72.0	173808	13	US-10-003-806-10	Sequence 10, Appl
247	14.4	72.0	366	17	US-10-085-783A-23201	Sequence 23201, A	320	14.4	72.0	176080	17	US-10-235-132A-43	Sequence 43, Appl
248	14.4	72.0	369	13	US-10-027-632-50282	Sequence 50282, A	321	14.4	72.0	201143	17	US-10-240-425-1099	Sequence 1099, Ap
249	14.4	72.0	369	17	US-10-027-632-50282	Sequence 50282, A	322	14.4	72.0	247544	18	US-10-322-696-55	Sequence 55, Appl
250	14.4	72.0	394	17	US-10-424-599-5269	Sequence 5269, App	323	14.4	72.0	268685	15	US-10-265-071-22	Sequence 22, Appl
251	14.4	72.0	404	18	US-10-723-860-1273	Sequence 1273, App	324	14.4	72.0	268685	15	US-10-025-966A-22	Sequence 22, Appl
252	14.4	72.0	416	9	US-09-770-423-464	Sequence 464, App	325	14.4	72.0	310122	18	US-10-417-375-1	Sequence 1, Appli
253	14.4	72.0	420	18	US-10-357-930-36525	Sequence 36525, A	326	14.4	72.0	561515	18	US-10-741-601-5682	Sequence 5682, Ap
254	14.4	72.0	443	17	US-10-242-535A-7168	Sequence 7168, App	327	14.4	72.0	561515	19	US-10-741-600-17730	Sequence 17730, A
255	14.4	72.0	443	17	US-10-085-783A-7168	Sequence 7168, App	328	14.2	71.0	25	15	US-10-098-263B-7945	Sequence 7945, Ap
256	14.4	72.0	445	9	US-09-960-352-8746	Sequence 8746, App	329	14.2	71.0	25	19	US-10-719-900-328444	Sequence 328444, A
257	14.4	72.0	470	10	US-09-918-995-28591	Sequence 28591, A	330	14.2	71.0	25	19	US-10-719-900-528634	Sequence 528634, A
258	14.4	72.0	472	18	US-10-357-930-36818	Sequence 36818, A	331	14.2	71.0	25	19	US-10-719-900-588448	Sequence 588448, A
259	14.4	72.0	490	10	US-09-918-995-903	Sequence 903, App	332	14.2	71.0	25	19	US-10-719-900-701534	Sequence 701534, A
260	14.4	72.0	512	13	US-10-027-632-267663	Sequence 267663, App	333	14.2	71.0	25	19	US-10-719-900-719032	Sequence 719032, A
261	14.4	72.0	512	17	US-10-027-632-267663	Sequence 267663, App	334	14.2	71.0	50	17	US-10-131-827-181	Sequence 181, App
262	14.4	72.0	520	18	US-10-357-930-6546	Sequence 6546, App	335	14.2	71.0	54	17	US-10-440-850-2263	Sequence 2263, App
263	14.4	72.0	524	16	US-10-425-115-53758	Sequence 53758, A	336	14.2	71.0	75	18	US-10-674-124A-22635	Sequence 22635, A
264	14.4	72.0	525	18	US-10-029-386-12883	Sequence 12883, A	337	14.2	71.0	93	18	US-10-674-124A-4176	Sequence 4176, Ap
265	14.4	72.0	543	13	US-10-027-632-90425	Sequence 90425, A	338	14.2	71.0	95	9	US-09-864-761-22011	Sequence 22011, Ap
266	14.4	72.0	543	17	US-10-027-632-90425	Sequence 90425, A	339	14.2	71.0	111	18	US-10-674-124A-18235	Sequence 18235, A
267	14.4	72.0	545	9	US-09-925-300-266	Sequence 266, App	340	14.2	71.0	129	18	US-10-674-124A-18749	Sequence 18749, A
268	14.4	72.0	547	16	US-10-029-386-768	Sequence 768, App	341	14.2	71.0	138	18	US-10-674-124A-19715	Sequence 19715, A
269	14.4	72.0	551	13	US-10-027-632-51580	Sequence 51580, A	342	14.2	71.0	138	18	US-10-674-124A-19715	Sequence 19715, A
270	14.4	72.0	551	13	US-10-027-632-53839	Sequence 53839, A	343	14.2	71.0	139	18	US-10-767-701-19286	Sequence 19286, A
271	14.4	72.0	551	13	US-10-027-632-64643	Sequence 64643, A	344	14.2	71.0	158	18	US-10-674-124A-18552	Sequence 18552, A
272	14.4	72.0	551	13	US-10-027-632-321516	Sequence 321516, A	345	14.2	71.0	160	18	US-10-674-124A-909	Sequence 909, App
273	14.4	72.0	551	17	US-10-027-632-51580	Sequence 51580, A	346	14.2	71.0	165	18	US-10-674-124A-910	Sequence 910, App
274	14.4	72.0	551	17	US-10-027-632-53839	Sequence 53839, A	347	14.2	71.0	169	18	US-10-425-115-6203	Sequence 6203, App
275	14.4	72.0	551	17	US-10-027-632-64643	Sequence 64643, A	348	14.2	71.0	177	17	US-10-243-552-26	Sequence 26, Appl
276	14.4	72.0	551	17	US-10-027-632-321516	Sequence 321516, A	349	14.2	71.0	201	18	US-10-437-963-77372	Sequence 77372, A
277	14.4	72.0	564	13	US-10-027-632-267662	Sequence 267662, A	350	14.2	71.0	201	18	US-10-741-601-7589	Sequence 7589, Ap
278	14.4	72.0	564	17	US-10-027-632-267662	Sequence 267662, A	351	14.2	71.0	201	18	US-10-741-601-7665	Sequence 7665, Ap
279	14.4	72.0	582	18	US-10-357-930-21137	Sequence 21137, A	352	14.2	71.0	201	18	US-10-741-601-7677	Sequence 7677, Ap
280	14.4	72.0	882	18	US-10-357-930-26976	Sequence 26976, A	353	14.2	71.0	201	18	US-10-719-993-3608	Sequence 3608, Ap
281	14.4	72.0	933	19	US-10-774-355A-141	Sequence 141, App	354	14.2	71.0	201	18	US-10-719-993-17812	Sequence 17812, A
282	14.4	72.0	1139	17	US-10-424-599-43911	Sequence 43911, A	355	14.2	71.0	201	18	US-10-719-993-29153	Sequence 29153, A
283	14.4	72.0	1237	10	US-09-984-276-66	Sequence 66, Appl	356	14.2	71.0	201	18	US-10-719-993-29704	Sequence 29704, A
284	14.4	72.0	1237	10	US-09-984-271-66	Sequence 66, Appl	357	14.2	71.0	201	18	US-10-719-993-42824	Sequence 42824, A
285	14.4	72.0	1340	18	US-10-357-930-24176	Sequence 24176, A	358	14.2	71.0	201	19	US-10-741-600-21382	Sequence 21382, A
286	14.4	72.0	1340	18	US-10-357-930-25496	Sequence 25496, A	359	14.2	71.0	201	19	US-10-741-600-21458	Sequence 21458, A
287	14.4	72.0	1412	18	US-10-357-930-24391	Sequence 24391, A	360	14.2	71.0	201	19	US-10-741-600-21460	Sequence 21460, A
288	14.4	72.0	1412	18	US-10-357-930-24486	Sequence 24486, A	361	14.2	71.0	201	19	US-10-741-600-21470	Sequence 21470, A
289	14.4	72.0	1740	17	US-10-108-260A-1833	Sequence 1833, App	362	14.2	71.0	216	10	US-09-918-995-32728	Sequence 32728, A
290	14.4	72.0	2043	9	US-09-974-300-1653	Sequence 1653, App	363	14.2	71.0	223	9	US-09-864-761-18011	Sequence 18011, A
291	14.4	72.0	2254	9	US-09-925-301-453	Sequence 453, App	364	14.2	71.0	223	18	US-10-425-115-24218	Sequence 24218, A
292	14.4	72.0	2393	13	US-10-098-841-47	Sequence 47, Appl	365	14.2	71.0	238	18	US-10-674-124A-12659	Sequence 12659, A
293	14.4	72.0	2483	17	US-10-104-047-252	Sequence 252, App	366	14.2	71.0	245	9	US-09-922-217-576	Sequence 576, App
294	14.4	72.0	2520	17	US-10-159-563-107	Sequence 107, App	367	14.2	71.0	245	9	US-09-833-263-576	Sequence 576, App
295	14.4	72.0	2530	18	US-10-723-860-5768	Sequence 5768, App	368	14.2	71.0	245	13	US-10-025-380-576	Sequence 576, App
296	14.4	72.0	2725	9	US-09-964-824A-243	Sequence 243, App	369	14.2	71.0	259	17	US-10-424-599-110379	Sequence 110379, A
297	14.4	72.0	2725	9	US-09-880-107-2181	Sequence 2181, App	370	14.2	71.0	279	18	US-10-425-115-11452	Sequence 11452, A
298	14.4	72.0	2861	18	US-10-437-963-36373	Sequence 36373, A	371	14.2	71.0	314	10	US-09-918-995-19016	Sequence 19016, A
299	14.4	72.0	3020	18	US-10-723-860-5474	Sequence 5474, App	372	14.2	71.0	321	18	US-10-425-115-20683	Sequence 20683, A
300	14.4	72.0	3846	14	US-10-198-846-10269	Sequence 10269, A	373	14.2	71.0	328	18	US-10-425-115-32214	Sequence 32214, A
301	14.4	72.0	4775	17	US-10-191-803-88	Sequence 88, Appl	374	14.2	71.0	335	9	US-09-864-761-22177	Sequence 22177, A
302	14.4	72.0	4775	17	US-10-152-319A-1657	Sequence 1657, App	375	14.2	71.0	340	18	US-10-674-124A-2741	Sequence 2741, Ap
303	14.4	72.0	5253	9	US-09-880-107-1655	Sequence 1655, App	376	14.2	71.0	345	17	US-10-242-535A-38239	Sequence 38239, A

C 377	14.2	71.0	345	17	US-10-085-783A-38239	Sequence 38239, A	C 450	14.2	71.0	663	17	US-10-027-632-234171	Sequence 234171, A
C 378	14.2	71.0	348	17	US-10-242-535A-11103	Sequence 11103, A	C 451	14.2	71.0	668	13	US-10-027-632-114166	Sequence 114166, A
C 379	14.2	71.0	348	17	US-10-085-783A-11103	Sequence 11103, A	C 452	14.2	71.0	668	13	US-10-027-632-114167	Sequence 114167, A
C 380	14.2	71.0	358	9	US-09-864-761-4530	Sequence 4530, Ap	C 453	14.2	71.0	688	17	US-10-027-632-114166	Sequence 114166, A
C 381	14.2	71.0	362	18	US-10-437-963-57653	Sequence 57653, A	C 454	14.2	71.0	688	17	US-10-027-632-114167	Sequence 114167, A
C 382	14.2	71.0	365	18	US-10-674-124A-14420	Sequence 14420, A	C 455	14.2	71.0	690	13	US-10-027-632-240712	Sequence 240712, A
C 383	14.2	71.0	367	18	US-10-425-115-123687	Sequence 123687, A	C 456	14.2	71.0	690	17	US-10-027-632-240712	Sequence 240712, A
C 384	14.2	71.0	389	18	US-10-674-124A-4515	Sequence 4515, Ap	C 457	14.2	71.0	707	14	US-10-175-523-73	Sequence 73, Appl
C 385	14.2	71.0	389	18	US-10-674-124A-25377	Sequence 25377, A	C 458	14.2	71.0	712	13	US-10-027-632-167143	Sequence 167143, A
C 386	14.2	71.0	396	17	US-10-242-535A-10014	Sequence 10014, A	C 459	14.2	71.0	712	17	US-10-027-632-167143	Sequence 167143, A
C 387	14.2	71.0	396	17	US-10-242-535A-39623	Sequence 39623, A	C 460	14.2	71.0	712	13	US-10-027-632-150816	Sequence 150816, A
C 388	14.2	71.0	396	17	US-10-085-783A-10014	Sequence 10014, A	C 461	14.2	71.0	713	17	US-10-027-632-150816	Sequence 150816, A
C 389	14.2	71.0	396	17	US-10-085-783A-39623	Sequence 39623, A	C 462	14.2	71.0	713	18	US-10-357-930-45655	Sequence 45655, A
C 390	14.2	71.0	398	13	US-10-027-632-96085	Sequence 96085, A	C 463	14.2	71.0	713	18	US-10-357-930-45892	Sequence 45892, A
C 391	14.2	71.0	398	13	US-10-027-632-96085	Sequence 96085, A	C 464	14.2	71.0	721	13	US-10-027-632-19774	Sequence 19774, A
C 392	14.2	71.0	403	14	US-10-198-846-3149	Sequence 3149, Ap	C 465	14.2	71.0	721	17	US-10-027-632-19774	Sequence 19774, A
C 393	14.2	71.0	408	9	US-09-867-701-5113	Sequence 5113, Ap	C 466	14.2	71.0	734	17	US-10-287-092-21	Sequence 21, Appl
C 394	14.2	71.0	423	9	US-09-864-761-5233	Sequence 5233, Ap	C 467	14.2	71.0	738	18	US-10-425-115-110024	Sequence 110024, A
C 395	14.2	71.0	432	17	US-10-424-599-68806	Sequence 68806, A	C 468	14.2	71.0	739	17	US-10-287-092-23	Sequence 23, Appl
C 396	14.2	71.0	436	9	US-09-954-456-33	Sequence 33, Appl	C 469	14.2	71.0	744	17	US-10-425-114-17023	Sequence 17023, A
C 397	14.2	71.0	436	9	US-09-954-456-672	Sequence 672, App	C 470	14.2	71.0	747	13	US-10-027-632-148613	Sequence 148613, A
C 398	14.2	71.0	441	18	US-10-357-930-16066	Sequence 16066, A	C 471	14.2	71.0	747	17	US-10-027-632-148613	Sequence 148613, A
C 399	14.2	71.0	455	17	US-10-424-599-124659	Sequence 124659, A	C 472	14.2	71.0	748	13	US-10-027-632-161496	Sequence 161496, A
C 400	14.2	71.0	463	10	US-09-918-995-22038	Sequence 22038, A	C 473	14.2	71.0	748	17	US-10-027-632-161496	Sequence 161496, A
C 401	14.2	71.0	464	18	US-10-425-115-160314	Sequence 160314, A	C 474	14.2	71.0	767	13	US-10-027-632-154445	Sequence 154445, A
C 402	14.2	71.0	467	9	US-09-864-761-1250	Sequence 1250, Ap	C 475	14.2	71.0	767	17	US-10-027-632-154445	Sequence 154445, A
C 403	14.2	71.0	467	17	US-10-424-599-111248	Sequence 111248, A	C 476	14.2	71.0	774	9	US-09-731-872-134	Sequence 134, App
C 404	14.2	71.0	469	18	US-10-674-124A-21949	Sequence 21949, A	C 477	14.2	71.0	774	10	US-09-876-997-134	Sequence 134, App
C 405	14.2	71.0	470	9	US-09-864-761-6577	Sequence 6577, Ap	C 478	14.2	71.0	776	13	US-10-027-632-167142	Sequence 167142, A
C 406	14.2	71.0	471	9	US-09-960-352-3869	Sequence 3869, Ap	C 479	14.2	71.0	776	17	US-10-027-632-167142	Sequence 167142, A
C 407	14.2	71.0	473	11	US-09-969-034-847	Sequence 847, App	C 480	14.2	71.0	788	17	US-10-425-114-18128	Sequence 18128, A
C 408	14.2	71.0	478	10	US-09-918-995-2758	Sequence 2758, App	C 481	14.2	71.0	801	13	US-10-027-632-127379	Sequence 127379, A
C 409	14.2	71.0	483	18	US-10-437-963-78928	Sequence 78928, A	C 482	14.2	71.0	801	17	US-10-027-632-127379	Sequence 127379, A
C 410	14.2	71.0	493	13	US-10-027-632-283703	Sequence 283703, A	C 483	14.2	71.0	807	13	US-10-027-632-165907	Sequence 165907, A
C 411	14.2	71.0	493	17	US-10-027-632-283703	Sequence 283703, A	C 484	14.2	71.0	807	13	US-10-027-632-165907	Sequence 165907, A
C 412	14.2	71.0	495	10	US-09-918-995-37888	Sequence 37888, A	C 485	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 413	14.2	71.0	495	10	US-09-918-995-21557	Sequence 21557, A	C 486	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 414	14.2	71.0	520	18	US-10-425-115-113658	Sequence 113658, A	C 487	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 415	14.2	71.0	524	18	US-10-357-930-15827	Sequence 15827, A	C 488	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 416	14.2	71.0	532	13	US-10-027-632-282909	Sequence 282909, A	C 489	14.2	71.0	815	13	US-10-027-632-127378	Sequence 127378, A
C 417	14.2	71.0	532	17	US-10-027-632-282909	Sequence 282909, A	C 490	14.2	71.0	815	17	US-10-027-632-127378	Sequence 127378, A
C 418	14.2	71.0	537	17	US-10-424-599-46076	Sequence 46076, A	C 491	14.2	71.0	819	13	US-10-027-632-164811	Sequence 164811, A
C 419	14.2	71.0	553	19	US-10-643-775-466	Sequence 466, App	C 492	14.2	71.0	819	17	US-10-027-632-164811	Sequence 164811, A
C 420	14.2	71.0	556	18	US-10-357-930-30160	Sequence 30160, A	C 493	14.2	71.0	823	18	US-10-425-115-176861	Sequence 176861, A
C 421	14.2	71.0	556	18	US-10-357-930-30344	Sequence 30344, A	C 494	14.2	71.0	824	17	US-10-425-114-18356	Sequence 18356, A
C 422	14.2	71.0	575	18	US-10-357-930-50074	Sequence 50074, A	C 495	14.2	71.0	826	13	US-10-027-632-155117	Sequence 155117, A
C 423	14.2	71.0	577	18	US-10-357-930-50142	Sequence 50142, A	C 496	14.2	71.0	826	17	US-10-027-632-155117	Sequence 155117, A
C 424	14.2	71.0	592	14	US-10-198-846-7919	Sequence 7919, Ap	C 497	14.2	71.0	836	13	US-10-027-632-159588	Sequence 159588, A
C 425	14.2	71.0	599	17	US-10-424-599-2472	Sequence 2472, Ap	C 498	14.2	71.0	836	13	US-10-027-632-159589	Sequence 159589, A
C 426	14.2	71.0	601	14	US-10-198-846-7978	Sequence 7978, Ap	C 499	14.2	71.0	836	13	US-10-027-632-159588	Sequence 159588, A
C 427	14.2	71.0	611	19	US-10-487-556-40	Sequence 40, Appl	C 500	14.2	71.0	836	17	US-10-027-632-159588	Sequence 159588, A
C 428	14.2	71.0	612	13	US-10-027-632-187176	Sequence 187176, A							
C 429	14.2	71.0	612	17	US-10-027-632-187176	Sequence 187176, A							
C 430	14.2	71.0	618	13	US-10-027-632-149994	Sequence 149994, A							
C 431	14.2	71.0	618	14	US-10-198-846-9013	Sequence 9013, Ap							
C 432	14.2	71.0	618	17	US-10-027-632-149994	Sequence 149994, A							
C 433	14.2	71.0	624	17	US-10-369-493-43481	Sequence 43481, A							
C 434	14.2	71.0	628	13	US-10-027-632-240010	Sequence 240010, A							
C 435	14.2	71.0	628	17	US-10-027-632-240010	Sequence 240010, A							
C 436	14.2	71.0	633	18	US-10-619-739-2051	Sequence 2051, Ap							
C 437	14.2	71.0	635	18	US-10-425-115-57413	Sequence 57413, A							
C 438	14.2	71.0	635	18	US-10-425-115-57413	Sequence 57413, A							
C 439	14.2	71.0	639	13	US-10-027-632-95409	Sequence 95409, A							
C 440	14.2	71.0	639	13	US-10-027-632-318713	Sequence 318713, A							
C 441	14.2	71.0	639	17	US-10-027-632-95409	Sequence 95409, A							
C 442	14.2	71.0	639	17	US-10-027-632-318713	Sequence 318713, A							
C 443	14.2	71.0	643	14	US-10-198-846-2100	Sequence 2100, Ap							
C 444	14.2	71.0	654	18	US-10-767-701-10777	Sequence 10777, A							
C 445	14.2	71.0	660	13	US-10-027-632-18941	Sequence 18941, A							
C 446	14.2	71.0	660	17	US-10-027-632-18941	Sequence 18941, A							
C 447	14.2	71.0	663	13	US-10-027-632-234170	Sequence 234170, A							
C 448	14.2	71.0	663	13	US-10-027-632-234171	Sequence 234171, A							
C 449	14.2	71.0	663	17	US-10-027-632-234170	Sequence 234170, A							

## ALIGNMENTS

## RESULT 1

US-09-851-871-255  
; Sequence 255, Application US/09851871  
; Publication No. US20030176374A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Clarence Frank  
; APPLICANT: Karras, James G.  
; APPLICANT: Vickers, Timothy A.  
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the  
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein  
; FILE REFERENCE: ISPH-0543  
; CURRENT APPLICATION NUMBER: US/09/851,871  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: PCT/US00/14471  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 09/326,186  
; PRIOR FILING DATE: 1999-06-04



; PRIOR APPLICATION NUMBER: 08/777,266  
; PRIOR FILING DATE: 1996-12-31  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 255  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-09-851-871-255

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
|||  
Db 1 CGTGTCTCTGTCTAGTCCC 20

## RESULT 2

US-10-444-206-255  
; Sequence 255, Application US/1044206  
; Publication No. US20040023917A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Clarence Frank  
; APPLICANT: Vickers, Timothy A.  
; APPLICANT: Karras, James G.  
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the  
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/444,206  
; CURRENT FILING DATE: 2003-05-23  
; PRIOR APPLICATION NUMBER: 09/851,871  
; PRIOR FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: PCT/US00/14471  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 09/326,186  
; PRIOR FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 08/777,266  
; PRIOR FILING DATE: 1996-12-31  
; NUMBER OF SEQ ID NOS: 444  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 255  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-444-206-255

Query Match 100.0%; Score 20; DB 17; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
|||  
Db 1 CGTGTCTCTGTCTAGTCCC 20

## RESULT 3

US-10-641-962-255  
; Sequence 255, Application US/10641962  
; Publication No. US20040235164A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett et al.  
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the  
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein  
; FILE REFERENCE: 30566/39578  
; CURRENT APPLICATION NUMBER: US/10/641,962  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 444

; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 255  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic primer  
US-10-641-962-255

Query Match 100.0%; Score 20; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
|||  
Db 1 CGTGTCTCTGTCTAGTCCC 20

## RESULT 4

US-10-858-500-623  
; Sequence 623, Application US/10858500  
; Publication No. US20050014257A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosanne M. Crooke  
; APPLICANT: Mark J. Graham  
; TITLE OF INVENTION: MODULATION OF C-REACTIVE PROTEIN EXPRESSION  
; FILE REFERENCE: BIOL0014US  
; CURRENT APPLICATION NUMBER: US/10/858,500  
; CURRENT FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 09/912,724  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: US 60/475,272  
; PRIOR FILING DATE: 2003-06-02  
; PRIOR APPLICATION NUMBER: US 60/540,042  
; PRIOR FILING DATE: 2004-01-28  
; NUMBER OF SEQ ID NOS: 627  
; SEQ ID NO 623  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Antisense Oligonucleotide  
US-10-858-500-623

Query Match 100.0%; Score 20; DB 19; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
|||  
Db 1 CGTGTCTCTGTCTAGTCCC 20

## RESULT 5

US-09-837-867A-39/c  
; Sequence 39, Application US/09837867A  
; Patent No. US20020098542A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharpe, Arlene H.  
; APPLICANT: Borriello, Francescopaulo  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; TITLE OF INVENTION: No. US20020098542A1e1 Forms of T Cell Costimulatory  
; TITLE OF INVENTION: Molecules and Uses Therefor  
; FILE REFERENCE: BWI-120CPADV  
; CURRENT APPLICATION NUMBER: US/09/837,867A  
; CURRENT FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: 08/205,697  
; PRIOR FILING DATE: 1994-03-02  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 124

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(124)
US-09-837-867A-39

Query Match      100.0%; Score 20; DB 9; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 6
US-09-962-969-39/c
; Sequence 39, Application US/09962969
; Publication No. US20030045703A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Sharpe
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon
; APPLICANT: Nadler, Lee
; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory
; FILE REFERENCE: BWI-120CPUS
; CURRENT APPLICATION NUMBER: US/10/643,768
; CURRENT FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: US/09/837,867
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(124)
US-10-643-768-39

TITLE OF INVENTION: No. US20030045703A1el Forms of T Cell Costimulatory
Molecules and Uses Therefor

NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/962,969
FILING DATE: 24-Sep-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/702,525
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: BWI-120CPUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 107..124
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-962-969-39

Query Match      100.0%; Score 20; DB 10; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38
```

```
RESULT 7
US-10-643-768-39/c
; Sequence 39, Application US/10643768
; Publication No. US20040192899A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory
; FILE REFERENCE: BWI-120CPADV
; CURRENT APPLICATION NUMBER: US/10/643,768
; CURRENT FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: US/09/837,867
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(124)
US-10-643-768-39

Query Match      100.0%; Score 20; DB 18; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 8
US-09-796-692-7754/c
; Sequence 7754, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
```

```
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (574)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7754

Query Match      100.0%; Score 20; DB 9; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTGCTGTGCTAGTCC 20
Db      63  CGTGTGCTGTGCTAGTCC 44

RESULT 9
US-10-040-862-7754/c
; Sequence 7754, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: unsure
; LOCATION: (574)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-7754

Query Match      100.0%; Score 20; DB 14; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTGCTGTGCTAGTCC 20
Db      63  CGTGTGCTGTGCTAGTCC 44

RESULT 10
US-10-057-475B-7754/c
; Sequence 7754, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(598)
; OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-7754

Query Match      100.0%; Score 20; DB 17; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTGCTGTGCTAGTCC 20
Db      63  CGTGTGCTGTGCTAGTCC 44
```

```
RESULT 11
US-10-154-884B-7754/c
; Sequence 7754, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,201
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(598)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-7754

Query Match 100.0%; Score 20; DB 17; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 63 CGTGTGCTGTGCTAGTCCC 44

RESULT 12
US-10-764-324-7754/c
; Sequence 7754, Application US/10764324
; Publication No. US2004017539A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-25/c
; Sequence 25, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-25

Query Match 100.0%; Score 20; DB 10; Length 1112;
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38
```

```
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (574)
; OTHER INFORMATION: n=A,T,C or G
US-10-764-324-7754

Query Match 100.0%; Score 20; DB 18; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 63 CGTGTGCTGTGCTAGTCCC 44

RESULT 13
US-09-441-411-25/c
; Sequence 25, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-25

Query Match 100.0%; Score 20; DB 10; Length 1112;
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38
```

RESULT 14  
US-10-762-128-25/c  
; Sequence 25, Application US/10762128  
; Publication No. US20040219161A1  
; GENERAL INFORMATION:  
; APPLICANT: Scholler, Nathalie B.  
; APPLICANT: Disis, Mary L.  
; APPLICANT: Hellstrom, Ingegerd  
; APPLICANT: Hellstrom, Karl Erik  
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
; FILE REFERENCE: 730033.409C1  
; CURRENT APPLICATION NUMBER: US/10/762,128  
; CURRENT FILING DATE: 2004-01-20  
; PRIOR APPLICATION NUMBER: US 09/441,411  
; PRIOR FILING DATE: 1999-11-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 1112  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-762-128-25

Query Match 100.0%; Score 20; DB 18; Length 1112;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20  
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 15  
US-08-592-711-3/c  
; Sequence 3, Application US/08592711  
; Publication No. US20020115214A1  
; GENERAL INFORMATION:  
; APPLICANT: June, Carl H.  
; APPLICANT: Thompson, Craig B.  
; APPLICANT: Nabel, Gary J.  
; APPLICANT: Gray, Gary S.  
; APPLICANT: Rennert, Paul D.  
; TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/592,711  
; FILING DATE: 26-JAN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,816  
; FILING DATE: 4-MAY-1995  
; APPLICATION NUMBER: US 08/403,253  
; FILING DATE: 10-MARCH-1995  
; APPLICATION NUMBER: US 08/253,964  
; FILING DATE: 3-JUNE-1994  
; APPLICATION NUMBER: US 08/073,223  
; FILING DATE: 4-JUNE-1993  
; APPLICATION NUMBER: US 08/200,947  
; FILING DATE: 23-FEB-1994  
; APPLICATION NUMBER: US 07/864,805  
; FILING DATE: 7-APR-1992  
; APPLICATION NUMBER: US 08/247,505

; FILING DATE: 23-MAY-1994  
; APPLICATION NUMBER: US 07/864,866  
; FILING DATE: 7-APR-1992  
; APPLICATION NUMBER: US 08/218,155  
; FILING DATE: 25-MAR-1994  
; APPLICATION NUMBER: US 07/864,807  
; FILING DATE: 7-APR-1992  
; APPLICATION NUMBER: US 07/902,467  
; FILING DATE: 16-JUNE-1992  
; APPLICATION NUMBER: US 07/275,433  
; FILING DATE: 23-NOV-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: RPI-002CP4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 107..1093  
US-08-592-711-3

Query Match 100.0%; Score 20; DB 8; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20  
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 16  
US-09-183-055-3/c  
; Sequence 3, Application US/09183055  
; Publication No. US20020076407A1  
; GENERAL INFORMATION:  
; APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.  
; APPLICANT: Gray, Gary S., Rennert, Paul D.  
; TITLE OF INVENTION: Methods For Selectively Stimulating  
; Proliferation Of T-Cells  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HALE AND DORR LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/183,055  
; FILING DATE: 30-Oct-1998  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/403,253  
; FILING DATE: March 10, 1995  
; APPLICATION NUMBER: US 08/253,964  
; FILING DATE: 3 JUNE 1994  
; APPLICATION NUMBER: US 08/073,223  
; FILING DATE: 4 JUNE 1993  
; APPLICATION NUMBER: US 08/200,947

```
;
; FILING DATE: 23 FEB 1994
; APPLICATION NUMBER: US 07/864,805
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/247,505
; FILING DATE: 23 MAY 1994
; APPLICATION NUMBER: US 07/864,866
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/218,155
; FILING DATE: 25 MAR 1994
; APPLICATION NUMBER: US 07/864,807
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 07/902,467
; FILING DATE: 16 JUNE 1992
; APPLICATION NUMBER: US 07/275,433
; FILING DATE: 23 NOV 1988
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Superko, Colleen
; REGISTRATION NUMBER: 39,850
; REFERENCE/DOCKET NUMBER: 36119-125US8
; TELEPHONE: (617) 526-6564
; TELEFAX: (617) 526-5000
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-183-055-3

Query Match 100.0%; Score 20; DB 9; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 17
US-09-425-762-1/c
; Sequence 1, Application US/09425762
; Publication No. US20020086414A1
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. 6605279e1 CTLA4/CD28 Ligands and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,762
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,744
;

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-004CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-425-762-1

Query Match 100.0%; Score 20; DB 9; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 18
US-09-837-867A-22/c
; Sequence 22, Application US/09837867A
; Patent No. US20020098542A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. US20020098542A1el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: BWI-120CPADV
; CURRENT APPLICATION NUMBER: US/09/837,867A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(1093)
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-837-867A-22

Query Match 100.0%; Score 20; DB 9; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 19
US-09-962-969-22/c
; Sequence 22, Application US/09962969
; Publication No. US20030045703A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Sharpe
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon
```



;/ Nadler, Lee  
/ TITLE OF INVENTION: No. US20030045703a1el Forms of T Cell Costimulatory  
/ Molecules and Uses Therefor  
/ NUMBER OF SEQUENCES: 65  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: LAHIVE & COCKFIELD  
/ STREET: 28 State Street  
/ CITY: Boston  
/ STATE: Massachusetts  
/ COUNTRY: USA  
/ ZIP: 02109-1875  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: ASCII Text  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/962,969  
/ FILING DATE: 24-Sep-2001  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 08/702,525  
/ FILING DATE: <Unknown>  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Mandragouras, Amy E.  
/ REGISTRATION NUMBER: 36,207  
/ REFERENCE/DOCKET NUMBER: BWI-120CPUS  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (617)227-7400  
/ TELEFAX: (617)227-5941  
/ INFORMATION FOR SEQ ID NO: 22:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1120 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: double  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: CDNA  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 107..1093  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-962-969-22  
  
Query Match 100.0%; Score 20; DB 10; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CGTGTCTCTGTGCTAGTCCC 20  
Db 57 CGTGTCTCTGTGCTAGTCCC 38  
  
RESULT 20  
US-09-350-202-3/c  
/ Sequence 3, Application US/09350202  
/ Publication No. US20030099643A1  
/ GENERAL INFORMATION:  
/ APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.  
/ APPLICANT: Gray, Gary S., Rennert, Paul D.  
/ TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells  
/ NUMBER OF SEQUENCES: 14  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: LAHIVE & COCKFIELD  
/ STREET: 28 State Street  
/ CITY: Boston  
/ STATE: Massachusetts  
/ COUNTRY: USA  
/ ZIP: 02109  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:

;/ APPLICATION NUMBER: US/09/350,202  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 08/403,253  
/ FILING DATE: March 10, 1995  
/ APPLICATION NUMBER: US 08/253,964  
/ FILING DATE: 3 JUNE 1994  
/ APPLICATION NUMBER: US 08/073,223  
/ FILING DATE: 4 JUNE 1993  
/ APPLICATION NUMBER: US 08/200,947  
/ FILING DATE: 23 FEB 1994  
/ APPLICATION NUMBER: US 07/864,805  
/ FILING DATE: 7 APR 1992  
/ APPLICATION NUMBER: US 08/247,505  
/ FILING DATE: 23 MAY 1994  
/ APPLICATION NUMBER: US 07/864,866  
/ FILING DATE: 7 APR 1992  
/ APPLICATION NUMBER: US 08/218,155  
/ FILING DATE: 25 MAR 1994  
/ APPLICATION NUMBER: US 07/864,807  
/ FILING DATE: 7 APR 1992  
/ APPLICATION NUMBER: US 07/902,467  
/ FILING DATE: 16 JUNE 1992  
/ APPLICATION NUMBER: US 07/275,433  
/ FILING DATE: 23 NOV 1988  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Mandragouras, Amy E.  
/ REGISTRATION NUMBER: 36,207  
/ REFERENCE/DOCKET NUMBER: RPI-002CP2  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (617) 227-7400  
/ TELEFAX: (617) 742-4214  
/ INFORMATION FOR SEQ ID NO: 3:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1120 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: CDNA  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 107..1093  
US-09-350-202-3  
  
Query Match 100.0%; Score 20; DB 10; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CGTGTCTCTGTGCTAGTCCC 20  
Db 57 CGTGTCTCTGTGCTAGTCCC 38  
  
RESULT 21  
US-10-390-330-3/c  
/ Sequence 3, Application US/10390330  
/ Publication No. US20040001829A1  
/ GENERAL INFORMATION:  
/ APPLICANT: June, Carl H.  
/ APPLICANT: Thompson, Craig B.  
/ APPLICANT: Nabel, Gary J.  
/ APPLICANT: Gray, Gary S.  
/ APPLICANT: Rennert, Paul D.  
/ TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation  
/ NUMBER OF SEQUENCES: 14  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: LAHIVE & COCKFIELD  
/ STREET: 60 State Street, Suite 510  
/ CITY: Boston  
/ STATE: Massachusetts  
/ COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/390,330  
FILING DATE: March 17, 2003  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/435,816A  
FILING DATE: May 4, 1995  
CLASSIFICATION:  
APPLICATION NUMBER: US 08/403,253  
FILING DATE: 10 MARCH 1995  
APPLICATION NUMBER: US 08/253,964  
FILING DATE: 3 JUNE 1994  
APPLICATION NUMBER: US 08/073,223  
FILING DATE: 4 JUNE 1993  
APPLICATION NUMBER: US 08/200,947  
FILING DATE: 23 FEB 1994  
APPLICATION NUMBER: US 07/864,805  
FILING DATE: 7 APR 1992  
APPLICATION NUMBER: US 08/247,505  
FILING DATE: 23 MAY 1994  
APPLICATION NUMBER: US 07/864,866  
FILING DATE: 7 APR 1992  
APPLICATION NUMBER: US 08/218,155  
FILING DATE: 25 MAR 1994  
APPLICATION NUMBER: US 07/864,807  
FILING DATE: 7 APR 1992  
APPLICATION NUMBER: US 07/902,467  
FILING DATE: 16 JUNE 1992  
APPLICATION NUMBER: US 07/275,433  
FILING DATE: 23 NOV 1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: RPI-002CP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1120 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 107..1093  
US-10-390-330-3

Query Match 100.0%; Score 20; DB 17; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGTGTGCTGTGCTAGTCCC 20  
|||||  
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 22  
US-10-643-768-22/c  
Sequence 22, Application US/10643768  
Publication No. US20040192899A1  
GENERAL INFORMATION:  
APPLICANT: Sharpe, Arlene H.  
APPLICANT: Borriello, Francescopaulo  
APPLICANT: Freeman, Gordon J.  
APPLICANT: Nadler, Lee M.

TITLE OF INVENTION: Novel Forms of T Cell Costimulatory  
TITLE OF INVENTION: Molecules and Uses Therefor  
FILE REFERENCE: BWI-120CPADV  
CURRENT APPLICATION NUMBER: US/10/643,768  
PRIOR FILING DATE: 2003-08-18  
PRIOR APPLICATION NUMBER: US/09/837,867  
PRIOR FILING DATE: 2001-04-17  
PRIOR APPLICATION NUMBER: 08/205,697  
PRIOR FILING DATE: 1994-03-02  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 1120  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (107) ... (1093)  
US-10-643-768-22

Query Match 100.0%; Score 20; DB 18; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGTGTGCTGTGCTAGTCCC 20  
|||||  
Db 57 CGTGTGCTGTGCTAGTCCC 38

## RESULT 23

US-10-429-079B-1/c  
Sequence 1, Application US/10429079B  
Publication No. US20040230051A1  
GENERAL INFORMATION:  
APPLICANT: Freeman, Gordon J.  
Nadler, Lee M.  
Gray, Gary S.

TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS FOR INHIBITING THE  
INTERACTIONS OF B7-1 AND B7-2 WITH THEIR NATURAL  
LIGANDS

NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street, 24th Floor  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/429,079B  
FILING DATE: 02-May-2003  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/425,762  
FILING DATE: 22-OCT-1999  
APPLICATION NUMBER: 08/479,744  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: 08/280,757  
FILING DATE: 26-JUL-1994  
APPLICATION NUMBER: 08/147,773  
FILING DATE: 03-NOV-1993  
APPLICATION NUMBER: 08/109,393  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: 08/101,624  
FILING DATE: 26-JUL-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: RPI-004C3DVCN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1120 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 107..1093  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-429-079B-1

Query Match 100.0%; Score 20; DB 18; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
DB 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 24  
US-10-444-206-296/c  
; Sequence 296, Application US/10444206  
; Publication No. US20040023917A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Clarence Frank  
; APPLICANT: Vickers, Timothy A.  
; APPLICANT: Karras, James G.  
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the  
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/444,206  
; CURRENT FILING DATE: 2003-05-23  
; PRIOR APPLICATION NUMBER: 09/851,871  
; PRIOR FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: PCT/US00/14471  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 09/326,186  
; PRIOR FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 08/777,266  
; PRIOR FILING DATE: 1996-12-31  
; NUMBER OF SEQ ID NOS: 444  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 296  
; LENGTH: 2781  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (117)...(1106)  
US-10-444-206-296

Query Match 100.0%; Score 20; DB 17; Length 2781;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
DB 67 CGTGTCTCTGTCTAGTCCC 48

RESULT 25  
US-10-756-783-5/c  
; Sequence 5, Application US/10756783  
; Publication No. US20040209363A1  
; GENERAL INFORMATION:  
; APPLICANT: WATTS, Tania

APPLICANT: BUKCZYNSKI, Jacob  
APPLICANT: WEN, Tao  
; TITLE OF INVENTION: Methods and compositions for modulating T cell activation and  
; TITLE OF INVENTION: uses thereof  
; FILE REFERENCE: 12912/2  
; CURRENT APPLICATION NUMBER: US/10/756,783  
; PRIOR FILING DATE: 2004-01-12  
; PRIOR APPLICATION NUMBER: US 60/304,430  
; PRIOR FILING DATE: 2001-07-12  
; PRIOR APPLICATION NUMBER: PCT/CA02/01033  
; PRIOR FILING DATE: 2002-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 2781  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (117)...(1106)  
US-10-756-783-5

Query Match 100.0%; Score 20; DB 18; Length 2781;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
DB 67 CGTGTCTCTGTCTAGTCCC 48

RESULT 26  
US-10-641-962-296/c  
; Sequence 296, Application US/10641962  
; Publication No. US20040235164A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett et al.  
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the  
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein  
; FILE REFERENCE: 30566/39578  
; CURRENT APPLICATION NUMBER: US/10/641,962  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 444  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 296  
; LENGTH: 2781  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (117)...(1106)  
US-10-641-962-296

Query Match 100.0%; Score 20; DB 18; Length 2781;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
DB 67 CGTGTCTCTGTCTAGTCCC 48

RESULT 27  
US-10-444-206-297/c  
; Sequence 297, Application US/10444206  
; Publication No. US20040023917A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Clarence Frank  
; APPLICANT: Vickers, Timothy A.  
; APPLICANT: Karras, James G.  
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the  
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein  
; FILE REFERENCE:

```
; CURRENT APPLICATION NUMBER: US/10/444,206
; CURRENT FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: 09/851,871
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: PCT/US00/14471
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 09/326,186
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 08/777,266
; PRIOR FILING DATE: 1996 12 31
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 297
; LENGTH: 68001
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-444-206-297

Query Match      100.0%; Score 20; DB 17; Length 68001;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTGCTGTGCTAGTCCC 20
        |||||
DB      1067 CGTGTGCTGTGCTAGTCCC 1048

RESULT 28
US-10-641-962-297/c
; Sequence 297, Application US/10641962
; Publication No. US20040235164A1
; GENERAL INFORMATION:
; APPLICANT: Bennett et al.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; FILE REFERENCE: 30566/39578
; CURRENT APPLICATION NUMBER: US/10/641,962
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 297
; LENGTH: 68001
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-641-962-297

Query Match      100.0%; Score 20; DB 18; Length 68001;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTGCTGTGCTAGTCCC 20
        |||||
DB      1067 CGTGTGCTGTGCTAGTCCC 1048

RESULT 29
US-09-796-692-7817/c
; Sequence 7817, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
```

```
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (526)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (535)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7817

Query Match      92.0%; Score 18.4; DB 9; Length 551;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CGTGTGCTGTGCTAGTCCC 20
        |||||
DB      41 CGTGTGCTGTGCTAGTCCC 22

RESULT 30
US-10-040-862-7817/c
; Sequence 7817, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
```

;  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7817  
; LENGTH: 551  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (526)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (535)  
; OTHER INFORMATION: n=A,T,C or G  
US-10-040-862-7817

Query Match 92.0%; Score 18.4; DB 14; Length 551;  
Best Local Similarity 95.0%; Pred. No. 21;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCCC 20  
||||| |||||||||  
Db 41 CGTGTGCTGTGCTAGTCCC 22

## RESULT 31

US-10-057-475B-7817/c  
; Sequence 7817, Application US/10057475B  
; Publication No. US2004002068A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Clapper, Jonathan David  
; APPLICANT: Wang, Aijun  
; APPLICANT: Ordonez, Nadia  
; APPLICANT: Carter, Lauren  
; APPLICANT: McNeill, Patricia Dianne  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-014402US  
; CURRENT APPLICATION NUMBER: US/10/057,475B  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7817  
; LENGTH: 551  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)---(551)  
; OTHER INFORMATION: n = g, a, c or t  
US-10-154-884B-7817

Query Match 92.0%; Score 18.4; DB 17; Length 551;

;  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 10979  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7817  
; LENGTH: 551  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)---(551)  
; OTHER INFORMATION: n = g, a, c or t  
US-10-057-475B-7817

Query Match 92.0%; Score 18.4; DB 17; Length 551;  
Best Local Similarity 95.0%; Pred. No. 21;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCCC 20  
||||| |||||||||  
Db 41 CGTGTGCTGTGCTAGTCCC 22

## RESULT 32

US-10-154-884B-7817/c  
; Sequence 7817, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 11290  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7817  
; LENGTH: 551  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)---(551)  
; OTHER INFORMATION: n = g, a, c or t  
US-10-154-884B-7817

Query Match 92.0%; Score 18.4; DB 17; Length 551;

```
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 41 CGTGTGCTGTGCTAGTCCC 22

RESULT 33
US-10-764-324-7817/c
; Sequence 7817, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (526)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (535)
; OTHER INFORMATION: n=A,T,C or G
US-10-764-324-7817

Query Match 92.0%; Score 18.4; DB 18; Length 551;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 41 CGTGTGCTGTGCTAGTCCC 22

RESULT 34
US-10-424-599-120635/c
; Sequence 120635, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1063
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 120635
; LENGTH: 215
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_79941C.1
US-10-424-599-120635

Query Match 87.0%; Score 17.4; DB 17; Length 215;
Best Local Similarity 94.7%; Pred. No. 65;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGTCTGTGCTAGTCCC 20
Db 144 GTGTGTCTGTGCTAGTCCC 126

RESULT 35
US-10-767-701-18097
; Sequence 18097, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 18097
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3478-006-P1-K1-C2
US-10-767-701-18097

Query Match 84.0%; Score 16.8; DB 18; Length 319;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 20 CGTGTCTGTGCAAGTCCC 39

RESULT 36
US-09-764-891-1063
; Sequence 1063, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1063
```



```
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (298)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (368)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (372)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-1063

Query Match      80.0%; Score 16; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 GTGTGTCGTGCTAGT 17
    |||||
Db   70 GTGTGTCGTGCTAGT 85

RESULT 37
US-09-728-444-574
; Sequence 574, Application US/09728444
; Patent No. US20020161207A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: Lex-0100-USA
; FILE REFERENCE: LEX-0100-USA
; CURRENT APPLICATION NUMBER: US/09/728,444
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,360
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 1206
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(515)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-444-574

Query Match      80.0%; Score 16; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 CGTGTGTCGTGCTAGT 16
    |||||
Db   384 CGTGTGTCGTGCTAGT 399

RESULT 38
US-10-087-192-613/c
; Sequence 613, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
```

```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 613
; LENGTH: 41434
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(41434)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-613

Query Match      80.0%; Score 16; DB 13; Length 41434;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 CGTGTGTCGTGCTAGT 16
    |||||
Db   5400 CGTGTGTCGTGCTAGT 5385

RESULT 39
US-10-719-900-719033
; Sequence 719033, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 719033
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-719033

Query Match      79.0%; Score 15.8; DB 19; Length 25;
Best Local Similarity 89.5%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  2 GTGTGTCGTGCTAGTCCC 20
    |||||
Db   1 GTGTGTCGTGCTAGTACC 19

RESULT 40
US-10-674-124A-23858/c
; Sequence 23858, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: GENETIC POLYMORPHISM MARKERS
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
```

```
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 23858
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: AC008576.6_117929
; FEATURE:
; OTHER INFORMATION: Located on chromosome 19
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 10025297
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 513
US-10-674-124A-23858
```

```
Query Match 79.0%; Score 15.8; DB 18; Length 293;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 GTGTGCTCTGTGCTAGTCCC 20
||||| ||||| ||||| |||||
Db 187 GTGTGCTCTGTGCTATATCCC 169
```

```
RESULT 41
US-10-276-774-1153
; Sequence 1153, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hysseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1153
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = a,t,c or g
US-10-276-774-1153
```

```
Query Match 79.0%; Score 15.8; DB 17; Length 439;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CGTGTGCTCTGTGCTAGTCC 19
||||| ||||| ||||| |||||
Db 39 CGTGTGCTCTGTGCGAGGCC 57
```

```
RESULT 42
US-10-027-632-290491
; Sequence 290491, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 290491
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-290491
```

```
Query Match 79.0%; Score 15.8; DB 13; Length 618;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 GTGTGCTCTGTGCTAGTCCC 20
||||| ||||| ||||| |||||
Db 403 GTGTTCGTGTCAGTCCC 421
```

```
RESULT 43
US-10-027-632-290492
; Sequence 290492, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 290492
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-290492
```

```
Query Match 79.0%; Score 15.8; DB 13; Length 618;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 GTGTGCTCTGTGCTAGTCCC 20
```



```
/ PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146229
; LENGTH: 757
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-146229

Query Match          79.0%; Score 15.8; DB 17; Length 757;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCC 19
Db 567 CGTGTGCTGTGCGAGGCC 585

RESULT 48
US-10-027-632-164600/c
; Sequence 164600, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164600
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164600

Query Match          79.0%; Score 15.8; DB 13; Length 804;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCC 19
Db 529 CGTGTGCTGTGCGAGGCC 511

RESULT 49
US-10-027-632-164601/c
; Sequence 164601, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
/ APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164601
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164601

Query Match          79.0%; Score 15.8; DB 13; Length 804;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCC 19
Db 529 CGTGTGCTGTGCGAGGCC 511

RESULT 50
US-10-027-632-164600/c
; Sequence 164600, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164600
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164600

Query Match          79.0%; Score 15.8; DB 17; Length 804;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 CGTGTCTGTGTCTAGTCC 19  
|||||  
Db 529 CGTGTCTGTGTGTGGAGGCC 511  
|||||

Search completed: March 15, 2005, 22:52:08  
Job time : 518 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 18:05:05 ; Search time 127 Seconds  
(without alignments)  
257.681 Million cell updates/sec

Title: US-09-980-953-256  
Perfect score: 20  
Sequence: 1 cgtgtctgtctgtctgtcc 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : Issued Patents NA: +  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	124	3	US-08-205-697A-39
C 2	20	100.0	124	3	US-08-702-525-39
C 3	20	100.0	124	4	US-09-837-867A-39
C 4	20	100.0	124	5	PCT-US95-02576-39
C 5	20	100.0	430	4	US-09-513-999C-425
C 6	20	100.0	1112	4	US-09-441-411-25
C 7	20	100.0	1120	2	US-08-456-104-1
C 8	20	100.0	1120	2	US-08-101-624-1
C 9	20	100.0	1120	3	US-08-479-744A-1
C 10	20	100.0	1120	3	US-08-280-757B-1
C 11	20	100.0	1120	3	US-08-205-697A-22
C 12	20	100.0	1120	3	US-08-702-525-22
C 13	20	100.0	1120	3	US-08-403-253A-3
C 14	20	100.0	1120	4	US-08-435-816A-3
C 15	20	100.0	1120	4	US-09-425-762-1
C 16	20	100.0	1120	4	US-09-837-867A-22
C 17	20	100.0	1120	4	US-09-206-132-1
C 18	20	100.0	1120	4	US-09-425-516-1
C 19	20	100.0	1120	5	PCT-US95-02576-22
C 20	16.4	82.0	264206	4	US-09-949-016-12731
C 21	16.4	82.0	264304	4	US-09-949-016-13249
C 22	16	80.0	368	4	US-09-513-999C-9101
C 23	15.8	79.0	2304	4	US-09-799-451-340
C 24	15.8	79.0	26007	4	US-09-949-016-17222
C 25	15.8	79.0	85850	4	US-09-949-016-13424
C 26	15.8	79.0	151295	4	US-09-949-016-14568
C 27	15.8	79.0	151295	4	US-09-949-016-14569

79.0	151295	4	US-09-949-016-14570	Sequence 14570, A
79.0	151295	4	US-09-949-016-14571	Sequence 14571, A
79.0	151295	4	US-09-949-016-14572	Sequence 14572, A
79.0	393753	4	US-09-949-016-14573	Sequence 14573, A
79.0	393753	4	US-09-949-016-14574	Sequence 14574, A
79.0	818128	4	US-09-949-016-14546	Sequence 14546, A
79.0	818128	4	US-09-949-016-14547	Sequence 14547, A
79.0	818128	4	US-09-949-016-14548	Sequence 14548, A
79.0	818128	4	US-09-949-016-14549	Sequence 14549, A
79.0	818128	4	US-09-949-016-14550	Sequence 14550, A
79.0	818128	4	US-09-949-016-14551	Sequence 14551, A
79.0	818128	4	US-09-949-016-14552	Sequence 14552, A
79.0	818128	4	US-09-949-016-14553	Sequence 14553, A
79.0	818128	4	US-09-949-016-14554	Sequence 14554, A
79.0	818128	4	US-09-949-016-14555	Sequence 14555, A
79.0	818128	4	US-09-949-016-14556	Sequence 14556, A
79.0	818128	4	US-09-949-016-14557	Sequence 14557, A
79.0	818128	4	US-09-949-016-14558	Sequence 14558, A
79.0	818128	4	US-09-949-016-14559	Sequence 14559, A
79.0	818128	4	US-09-949-016-14560	Sequence 14560, A
79.0	818128	4	US-09-949-016-14561	Sequence 14561, A
79.0	818128	4	US-09-949-016-14562	Sequence 14562, A
79.0	818128	4	US-09-949-016-14564	Sequence 14564, A
79.0	818128	4	US-09-949-016-14565	Sequence 14565, A
79.0	818128	4	US-09-949-016-14566	Sequence 14566, A
79.0	818128	4	US-09-949-016-14567	Sequence 14567, A
79.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
79.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
77.0	229354	4	US-09-705-400-64	Sequence 64, Appli
76.0	1158	4	US-09-142-027A-11	Sequence 11, Appli
76.0	1158	4	US-09-142-027A-13	Sequence 13, Appli
76.0	1573	1	US-08-597-545-4	Sequence 4, Appli
76.0	1573	1	US-08-457-135-4	Sequence 4, Appli
76.0	1913	4	US-09-270-767-14086	Sequence 14086, A
76.0	6089	4	US-09-724-128A-5	Sequence 5, Appli
76.0	30856	4	US-09-949-016-14613	Sequence 14613, A
76.0	57726	4	US-09-949-016-16430	Sequence 16430, A
76.0	61178	4	US-09-949-016-17369	Sequence 17369, A
76.0	63475	4	US-09-949-016-14516	Sequence 14516, A
76.0	203475	4	US-09-949-016-14517	Sequence 14517, A
76.0	203475	4	US-09-949-016-14518	Sequence 14518, A
76.0	203475	4	US-09-949-016-14519	Sequence 14519, A
76.0	203475	4	US-09-949-016-17226	Sequence 17226, A
76.0	203475	4	US-09-949-016-17227	Sequence 17227, A
76.0	203475	4	US-09-949-016-17228	Sequence 17228, A
76.0	203475	4	US-09-949-016-17229	Sequence 17229, A
75.0	100	6	5436136-8	Patent No. 5436136
75.0	100	6	5436136-8	Patent No. 5436136
75.0	291	4	US-09-513-999C-33244	Sequence 33244, A
75.0	368	6	5436136-9	Patent No. 5436136
75.0	368	6	5436136-9	Patent No. 5436136
75.0	627	6	5436136-1	Patent No. 5436136
75.0	627	6	5436136-1	Patent No. 5436136
75.0	1566	2	US-08-875-802-1	Sequence 1, Appli
75.0	1566	3	US-09-305-723C-1	Sequence 1, Appli
74.0	377	4	US-09-270-767-3061	Sequence 3061, Ap
74.0	377	4	US-09-270-767-18343	Sequence 18343, A
74.0	408	4	US-09-328-352-969	Sequence 969, App
74.0	601	4	US-09-465-559-21	Sequence 21, Appli
74.0	601	4	US-09-949-016-21673	Sequence 21673, A
74.0	601	4	US-09-949-016-21674	Sequence 21674, A
74.0	601	4	US-09-949-016-61024	Sequence 61024, A
74.0	601	4	US-09-949-016-61025	Sequence 61025, A
74.0	601	4	US-09-949-016-79866	Sequence 79866, A
74.0	601	4	US-09-949-016-79867	Sequence 79867, A
74.0	601	4	US-09-949-016-156745	Sequence 156745, A
74.0	601	4	US-09-949-016-156746	Sequence 156746, A
74.0	705	4	US-09-465-559-3	Sequence 3, Appli
74.0	8948	4	US-09-949-016-14854	Sequence 14854, A
74.0	9364	4	US-09-949-016-15965	Sequence 15965, A
74.0	22497	4	US-09-949-016-13511	Sequence 13511, A
74.0	22965	4	US-09-949-016-15862	Sequence 15862, A
74.0	32414	4	US-09-949-016-16443	Sequence 16443, A

c 101	14.8	74.0	40429	4	US-08-311-731A-125	Sequence 125, App	174	14.2	71.0	601	4	US-09-949-016-29095	Sequence 29095, A
c 102	14.8	74.0	90472	4	US-09-949-016-14038	Sequence 14038, A	175	14.2	71.0	601	4	US-09-949-016-29638	Sequence 29638, A
c 103	14.8	74.0	101810	4	US-09-949-016-13180	Sequence 13180, A	176	14.2	71.0	601	4	US-09-949-016-29639	Sequence 29639, A
c 104	14.8	74.0	116425	4	US-09-949-016-11819	Sequence 11809, A	c 177	14.2	71.0	601	4	US-09-949-016-69342	Sequence 69342, A
c 105	14.8	74.0	193303	4	US-09-949-016-13180	Sequence 37, App	c 178	14.2	71.0	601	4	US-09-949-016-69342	Sequence 69342, A
c 106	14.8	74.0	193303	4	US-09-949-016-13180	Sequence 44, App	c 179	14.2	71.0	601	4	US-09-949-016-69342	Sequence 69342, A
c 107	14.8	74.0	231129	4	US-09-949-016-16110	Sequence 16110, A	c 180	14.2	71.0	601	4	US-09-949-016-120996	Sequence 120996, A
c 108	14.8	74.0	260247	4	US-09-949-016-13158	Sequence 13158, A	c 181	14.2	71.0	601	4	US-09-949-016-121040	Sequence 121040, A
c 109	14.8	74.0	268293	4	US-09-949-016-11934	Sequence 11934, A	c 182	14.2	71.0	601	4	US-09-949-016-121041	Sequence 121041, A
c 110	14.8	74.0	301828	4	US-09-949-016-13969	Sequence 13969, A	c 183	14.2	71.0	601	4	US-09-949-016-131309	Sequence 131309, A
c 111	14.4	72.0	441	4	US-09-621-976-13445	Sequence 13445, A	c 184	14.2	71.0	601	4	US-09-949-016-136302	Sequence 136302, A
c 112	14.4	72.0	462	4	US-09-621-976-13445	Sequence 14252, A	c 185	14.2	71.0	601	4	US-09-949-016-137230	Sequence 137230, A
c 113	14.4	72.0	601	4	US-09-949-016-104758	Sequence 104758, A	c 186	14.2	71.0	601	4	US-09-949-016-141937	Sequence 141937, A
c 114	14.4	72.0	601	4	US-09-949-016-104762	Sequence 104762, A	c 187	14.2	71.0	601	4	US-09-949-016-142175	Sequence 142175, A
c 115	14.4	72.0	1237	4	US-09-482-273-66	Sequence 66, App	c 188	14.2	71.0	601	4	US-09-949-016-142176	Sequence 142176, A
c 116	14.4	72.0	1639	4	US-09-976-594-523	Sequence 523, App	c 189	14.2	71.0	601	4	US-09-949-016-156428	Sequence 156428, A
c 117	14.4	72.0	2658	4	US-09-543-681A-235	Sequence 235, App	c 190	14.2	71.0	601	4	US-09-949-016-181251	Sequence 181251, A
c 118	14.4	72.0	3182	4	US-08-188-582-12	Sequence 12, App	c 191	14.2	71.0	783	4	US-09-252-991A-15414	Sequence 15414, A
c 119	14.4	72.0	3182	4	US-08-646-715-12	Sequence 12, App	c 192	14.2	71.0	801	4	US-09-252-991A-15414	Sequence 15414, A
c 120	14.4	72.0	4872	4	US-09-949-016-1526	Sequence 5526, App	c 193	14.2	71.0	954	4	US-09-393-634-61	Sequence 15476, A
c 121	14.4	72.0	6174	3	US-09-298-731-47	Sequence 47, App	c 194	14.2	71.0	1331	4	US-09-023-655-579	Sequence 579, App
c 122	14.4	72.0	9364	4	US-09-949-016-15965	Sequence 15965, A	c 195	14.2	71.0	1556	1	US-08-587-545-3	Sequence 3, App
c 123	14.4	72.0	10571	4	US-09-949-016-12885	Sequence 12885, A	c 196	14.2	71.0	1556	1	US-08-587-545-3	Sequence 3, App
c 124	14.4	72.0	14747	3	US-09-608-285A-42	Sequence 42, App	c 197	14.2	71.0	1634	3	US-09-149-476-291	Sequence 291, App
c 125	14.4	72.0	14747	4	US-09-557-800C-42	Sequence 42, App	c 198	14.2	71.0	1820	4	US-09-799-451-885	Sequence 885, App
c 126	14.4	72.0	15977	3	US-09-608-285A-59	Sequence 59, App	c 199	14.2	71.0	1827	4	US-09-786-681A-3	Sequence 3, App
c 127	14.4	72.0	16002	4	US-09-949-016-12057	Sequence 12057, A	c 200	14.2	71.0	2005	4	US-09-949-016-2038	Sequence 2038, App
c 128	14.4	72.0	16003	4	US-09-949-016-17194	Sequence 17194, A	c 201	14.2	71.0	2031	3	US-09-149-476-155	Sequence 155, App
c 129	14.4	72.0	16347	4	US-09-949-016-14645	Sequence 14645, A	c 202	14.2	71.0	2072	4	US-09-786-681A-1	Sequence 1, App
c 130	14.4	72.0	16347	4	US-09-949-016-14645	Sequence 14646, A	c 203	14.2	71.0	2194	4	US-09-620-312D-795	Sequence 795, App
c 131	14.4	72.0	18864	4	US-09-949-016-16992	Sequence 16992, A	c 204	14.2	71.0	2358	4	US-09-919-039-190	Sequence 190, App
c 132	14.4	72.0	18864	4	US-09-949-016-16992	Sequence 16993, A	c 205	14.2	71.0	2538	4	US-09-023-655-25	Sequence 25, App
c 133	14.4	72.0	20840	4	US-09-949-016-14115	Sequence 14115, A	c 206	14.2	71.0	3190	4	US-09-710-279-3636	Sequence 3636, App
c 134	14.4	72.0	32666	4	US-09-949-016-16086	Sequence 16086, A	c 207	14.2	71.0	3364	4	US-09-710-279-3636	Sequence 3636, App
c 135	14.4	72.0	53442	4	US-09-949-016-11921	Sequence 11921, A	c 208	14.2	71.0	3501	4	US-09-710-279-3346	Sequence 3346, App
c 136	14.4	72.0	53453	4	US-09-949-016-13370	Sequence 13370, A	c 209	14.2	71.0	4117	3	US-08-984-970B-2	Sequence 2, App
c 137	14.4	72.0	66480	4	US-09-949-016-14817	Sequence 14817, A	c 210	14.2	71.0	4985	3	US-08-998-416-7	Sequence 7, App
c 138	14.4	72.0	87644	4	US-09-949-016-16041	Sequence 16041, A	c 211	14.2	71.0	5379	4	US-09-949-016-13457	Sequence 13457, A
c 139	14.4	72.0	97376	4	US-09-949-016-16093	Sequence 16093, A	c 212	14.2	71.0	5659	4	US-09-321-987B-1	Sequence 1, App
c 140	14.4	72.0	119032	4	US-09-949-016-12160	Sequence 12160, A	c 213	14.2	71.0	9127	4	US-09-949-016-13535	Sequence 13535, A
c 141	14.4	72.0	119032	4	US-09-949-016-17268	Sequence 17268, A	c 214	14.2	71.0	9544	4	US-09-949-016-16367	Sequence 16367, A
c 142	14.4	72.0	129415	4	US-09-949-016-16997	Sequence 16997, A	c 215	14.2	71.0	9885	4	US-09-949-016-12934	Sequence 12934, A
c 143	14.4	72.0	178122	4	US-09-949-016-16723	Sequence 16723, A	c 216	14.2	71.0	9885	4	US-09-949-016-12935	Sequence 12935, A
c 144	14.4	72.0	374159	4	US-09-949-016-15868	Sequence 15868, A	c 217	14.2	71.0	10367	1	US-08-110-300A-9	Sequence 9, App
c 145	14.4	72.0	462589	4	US-09-949-016-13900	Sequence 12900, A	c 218	14.2	71.0	10367	2	US-08-896-642-9	Sequence 9, App
c 146	14.4	72.0	476044	4	US-09-949-016-12412	Sequence 12412, A	c 219	14.2	71.0	10367	5	PCT-US93-08041-9	Sequence 9, App
c 147	14.4	72.0	767677	4	US-09-949-016-12147	Sequence 12147, A	c 220	14.2	71.0	11133	4	US-09-949-016-16567	Sequence 16567, A
c 148	14.4	72.0	767677	4	US-09-949-016-17361	Sequence 17361, A	c 221	14.2	71.0	12665	4	US-09-949-016-16388	Sequence 16388, A
c 149	14.2	71.0	54	2	US-08-585-684B-2701	Sequence 2701, App	c 222	14.2	71.0	13827	4	US-09-949-016-15599	Sequence 15599, A
c 150	14.2	71.0	54	3	US-09-038-073-2701	Sequence 2701, App	c 223	14.2	71.0	14634	4	US-09-949-016-17384	Sequence 17384, A
c 151	14.2	71.0	133	4	US-09-513-999C-9899	Sequence 9899, App	c 224	14.2	71.0	17284	4	US-09-949-016-12267	Sequence 12267, A
c 152	14.2	71.0	159	4	US-09-471-276-143	Sequence 143, App	c 225	14.2	71.0	17284	4	US-09-949-016-16000	Sequence 16000, A
c 153	14.2	71.0	219	4	US-09-513-999C-25457	Sequence 25457, A	c 226	14.2	71.0	18075	4	US-09-949-016-16643	Sequence 16643, A
c 154	14.2	71.0	238	4	US-09-513-999C-17226	Sequence 17226, A	c 227	14.2	71.0	19826	4	US-09-949-016-16793	Sequence 16793, A
c 155	14.2	71.0	432	4	US-09-513-999C-8272	Sequence 8272, App	c 228	14.2	71.0	25052	4	US-09-949-016-14585	Sequence 14585, A
c 156	14.2	71.0	446	4	US-09-270-767-11052	Sequence 11052, A	c 229	14.2	71.0	27632	4	US-09-949-016-12903	Sequence 12903, A
c 157	14.2	71.0	486	4	US-09-252-991A-15526	Sequence 15526, A	c 230	14.2	71.0	28030	4	US-09-949-016-12305	Sequence 12305, A
c 158	14.2	71.0	529	4	US-09-621-976-16089	Sequence 16089, A	c 231	14.2	71.0	28030	4	US-09-949-016-16948	Sequence 16948, A
c 159	14.2	71.0	576	1	US-08-086-428B-31	Sequence 31, App	c 232	14.2	71.0	28198	4	US-09-949-016-12348	Sequence 12348, A
c 160	14.2	71.0	576	1	US-08-086-428B-32	Sequence 32, App	c 233	14.2	71.0	32488	4	US-09-949-016-15490	Sequence 15490, A
c 161	14.2	71.0	576	1	US-08-086-428B-33	Sequence 33, App	c 234	14.2	71.0	35060	3	US-08-814-095-7	Sequence 7, App
c 162	14.2	71.0	576	2	US-08-468-570-31	Sequence 31, App	c 235	14.2	71.0	42395	4	US-09-949-016-12752	Sequence 12752, A
c 163	14.2	71.0	576	2	US-08-468-570-32	Sequence 32, App	c 236	14.2	71.0	4395	4	US-09-949-016-15115	Sequence 15115, A
c 164	14.2	71.0	576	2	US-08-468-570-33	Sequence 33, App	c 237	14.2	71.0	43069	4	US-09-252-542A-1	Sequence 1, App
c 165	14.2	71.0	576	2	US-08-290-665A-31	Sequence 31, App	c 238	14.2	71.0	43366	4	US-09-949-016-14185	Sequence 14185, A
c 166	14.2	71.0	576	2	US-08-290-665A-32	Sequence 32, App	c 239	14.2	71.0	47471	4	US-09-949-016-12271	Sequence 12271, A
c 167	14.2	71.0	576	2	US-08-290-665A-33	Sequence 33, App	c 240	14.2	71.0	48794	4	US-09-949-016-15637	Sequence 15637, A
c 168	14.2	71.0	576	4	US-08-466-601A-31	Sequence 31, App	c 241	14.2	71.0	49399	4	US-09-949-016-13780	Sequence 13780, A
c 169	14.2	71.0	576	4	US-08-466-601A-32	Sequence 32, App	c 242	14.2	71.0	51049	4	US-09-949-016-15571	Sequence 15571, A
c 170	14.2	71.0	576	4	US-08-466-601A-33	Sequence 33, App	c 243	14.2	71.0	63183	4	US-09-949-016-13047	Sequence 13047, A
c 171	14.2	71.0	576	5	PCT-US95-10398-31	Sequence 31, App	c 244	14.2	71.0	63183	4	US-09-949-016-13048	Sequence 13048, A
c 172	14.2	71.0	576	5	PCT-US95-10398-32	Sequence 32, App	c 245	14.2	71.0	64309	4	US-09-949-016-14581	Sequence 14581, A
c 173	14.2	71.0	576	5	PCT-US95-10398-33	Sequence 33, App	c 246	14.2	71.0	73853	4	US-09-949-016-12029	Sequence 12029, A



c 393 13.8 69.0 58768 4 US-09-949-016-13175 Sequence 13175, A  
394 13.8 69.0 59719 4 US-09-949-016-15616 Sequence 15616, A  
395 13.8 69.0 61178 4 US-09-949-016-17369 Sequence 17369, A  
396 13.8 69.0 61664 4 US-09-949-016-13308 Sequence 13308, A  
397 13.8 69.0 70308 4 US-09-949-016-15601 Sequence 15601, A  
398 13.8 69.0 75929 4 US-09-949-016-15543 Sequence 15543, A  
399 13.8 69.0 75929 4 US-09-949-016-15544 Sequence 15544, A  
400 13.8 69.0 77626 4 US-09-949-016-12608 Sequence 12608, A  
401 13.8 69.0 83617 4 US-09-949-016-12254 Sequence 12254, A  
402 13.8 69.0 84839 4 US-09-949-016-15816 Sequence 15816, A  
403 13.8 69.0 87190 4 US-09-949-016-16335 Sequence 16335, A  
c 404 13.8 69.0 89210 4 US-09-949-016-16988 Sequence 16988, A  
c 405 13.8 69.0 91665 4 US-09-949-016-12234 Sequence 12234, A  
406 13.8 69.0 95621 4 US-09-949-016-13237 Sequence 13237, A  
407 13.8 69.0 97195 4 US-09-949-016-12212 Sequence 12212, A  
408 13.8 69.0 97196 4 US-09-949-016-16971 Sequence 16971, A  
409 13.8 69.0 100928 4 US-09-949-016-16926 Sequence 16926, A  
410 13.8 69.0 113100 4 US-09-949-016-12245 Sequence 12245, A  
411 13.8 69.0 113876 4 US-09-949-016-14828 Sequence 14828, A  
412 13.8 69.0 113876 4 US-09-949-016-14829 Sequence 14829, A  
413 13.8 69.0 115508 4 US-09-949-016-11800 Sequence 11800, A  
414 13.8 69.0 115508 4 US-09-949-016-14826 Sequence 14826, A  
c 415 13.8 69.0 115508 4 US-09-949-016-14827 Sequence 14827, A  
c 416 13.8 69.0 133063 4 US-09-949-016-16293 Sequence 16293, A  
417 13.8 69.0 155266 4 US-09-949-016-13870 Sequence 13870, A  
418 13.8 69.0 161124 4 US-09-949-016-11760 Sequence 11760, A  
419 13.8 69.0 175236 4 US-09-949-016-14353 Sequence 14353, A  
420 13.8 69.0 211049 4 US-09-949-016-15770 Sequence 15770, A  
c 421 13.8 69.0 211049 4 US-09-949-016-15770 Sequence 15770, A  
422 13.8 69.0 236474 4 US-09-949-016-13418 Sequence 13418, A  
423 13.8 69.0 283538 4 US-09-949-016-13506 Sequence 13506, A  
424 13.8 69.0 298336 4 US-09-949-016-16600 Sequence 16600, A  
425 13.8 69.0 312957 4 US-09-949-001-31 Sequence 31, Appl  
426 13.8 69.0 312972 4 US-09-949-001-34 Sequence 34, Appl  
427 13.8 69.0 343352 4 US-09-949-016-13498 Sequence 13498, A  
428 13.8 69.0 392000 4 US-10-027-983-11 Sequence 11, Appl  
429 13.8 69.0 450395 4 US-09-949-016-15473 Sequence 15473, A  
c 430 13.6 68.0 106 3 US-08-482-677-13 Sequence 13, Appl  
c 431 13.6 68.0 106 4 US-10-033-174-13 Sequence 13, Appl  
432 13.6 68.0 187 4 US-09-270-767-3794 Sequence 3794, Ap  
433 13.6 68.0 187 4 US-09-270-767-19076 Sequence 19076, A  
c 434 13.6 68.0 231 4 US-09-902-540-1610 Sequence 1610, Ap  
435 13.6 68.0 310 4 US-09-513-999C-30296 Sequence 30296, A  
c 436 13.6 68.0 329 4 US-09-621-976-7848 Sequence 7848, Ap  
437 13.6 68.0 331 4 US-09-513-999C-32221 Sequence 32221, A  
438 13.6 68.0 367 3 US-08-908-223-180 Sequence 180, App  
439 13.6 68.0 415 4 US-09-270-767-29304 Sequence 29304, A  
c 440 13.6 68.0 416 4 US-09-621-976-136 Sequence 136, App  
c 441 13.6 68.0 443 4 US-09-621-976-14153 Sequence 14153, A  
442 13.6 68.0 443 4 US-09-270-767-3949 Sequence 3949, Ap  
443 13.6 68.0 443 4 US-09-270-767-19231 Sequence 19231, A  
c 444 13.6 68.0 566 4 US-09-949-016-83728 Sequence 83728, A  
c 445 13.6 68.0 601 4 US-09-949-016-22192 Sequence 22192, A  
c 446 13.6 68.0 601 4 US-09-949-016-22193 Sequence 22193, A  
c 447 13.6 68.0 601 4 US-09-949-016-26341 Sequence 26341, A  
c 448 13.6 68.0 601 4 US-09-949-016-26342 Sequence 26342, A  
c 449 13.6 68.0 601 4 US-09-949-016-68800 Sequence 68800, A  
c 450 13.6 68.0 601 4 US-09-949-016-78595 Sequence 78595, A  
c 451 13.6 68.0 601 4 US-09-949-016-83729 Sequence 83729, A  
c 452 13.6 68.0 601 4 US-09-949-016-83730 Sequence 83730, A  
c 453 13.6 68.0 601 4 US-09-949-016-91772 Sequence 91772, A  
454 13.6 68.0 601 4 US-09-949-016-105720 Sequence 105720, A  
c 455 13.6 68.0 601 4 US-09-949-016-105728 Sequence 105728, A  
c 456 13.6 68.0 601 4 US-09-949-016-130368 Sequence 130368, A  
457 13.6 68.0 601 4 US-09-949-016-140995 Sequence 140995, A  
458 13.6 68.0 601 4 US-09-949-016-140996 Sequence 140996, A  
c 459 13.6 68.0 601 4 US-09-949-016-156013 Sequence 156013, A  
c 460 13.6 68.0 601 4 US-09-949-016-170283 Sequence 170283, A  
461 13.6 68.0 601 4 US-09-949-016-176149 Sequence 176149, A  
c 462 13.6 68.0 601 4 US-09-949-016-189048 Sequence 189048, A  
c 463 13.6 68.0 601 4 US-09-949-016-189049 Sequence 189049, A  
c 464 13.6 68.0 601 4 US-09-949-016-194713 Sequence 194713, A  
465 13.6 68.0 601 4 US-09-949-016-203029 Sequence 203029, A

Sequence 203037, App  
Sequence 205849, App  
Sequence 205850, App  
Sequence 109, App  
Sequence 628, App  
Sequence 232, App  
Sequence 157, App  
Sequence 158, App  
Sequence 13352, A  
Sequence 4554, Ap  
Sequence 5, Appli  
Sequence 236, App  
Sequence 138, App  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 7089, Ap  
Sequence 9, Appli  
Sequence 760, App  
Sequence 1, Appli  
Sequence 7, Appli  
Sequence 1, Appli  
Sequence 15, Appl  
Sequence 13, Appl  
Sequence 253, App  
Sequence 5522, App  
Sequence 713, App  
Sequence 38, Appl  
Sequence 38, Appl  
Sequence 246, App  
Sequence 11, Appl  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 5809, Ap  
Sequence 207, App

13.6 68.0 601 4 US-09-949-016-203037  
13.6 68.0 601 4 US-09-949-016-205849  
13.6 68.0 601 4 US-09-949-016-205850  
13.6 68.0 716 4 US-09-023-655-109  
13.6 68.0 720 3 US-08-998-416-628  
13.6 68.0 905 4 US-09-976-594-232  
13.6 68.0 1001 4 US-09-671-317-157  
13.6 68.0 1001 4 US-09-671-317-158  
13.6 68.0 1007 4 US-09-270-767-13352  
13.6 68.0 1057 4 US-09-949-016-4554  
13.6 68.0 1119 4 US-09-784-358-5  
13.6 68.0 1244 4 US-09-902-540-236  
13.6 68.0 1289 3 US-09-247-155-138  
13.6 68.0 1341 4 US-09-784-358-3  
13.6 68.0 1430 3 US-09-492-985-1  
13.6 68.0 1797 4 US-09-489-039A-7089  
13.6 68.0 1953 4 US-09-784-358-9  
13.6 68.0 1969 4 US-09-799-451-760  
13.6 68.0 2005 1 US-08-208-108-1  
13.6 68.0 2175 4 US-09-784-358-7  
13.6 68.0 2181 3 US-09-126-980-1  
13.6 68.0 2181 3 US-09-476-482-1  
13.6 68.0 2181 3 US-09-517-605-15  
13.6 68.0 2316 4 US-09-784-358-13  
13.6 68.0 2332 4 US-09-949-016-253  
13.6 68.0 2332 4 US-09-949-016-5522  
13.6 68.0 2425 4 US-09-799-451-713  
13.6 68.0 2425 4 US-09-894-988A-38  
13.6 68.0 2517 4 US-10-237-551-38  
13.6 68.0 2517 4 US-10-237-551-246  
13.6 68.0 2538 4 US-09-784-358-11  
13.6 68.0 2931 4 US-09-623-624-1  
13.6 68.0 2931 4 US-10-270-595-1  
13.6 68.0 2934 4 US-09-949-016-5809  
13.6 68.0 2943 4 US-09-949-016-207

ALIGNMENTS

RESULT 1  
US-08-205-697A-39/c  
; Sequence 39, Application US/08205697A  
; Patent No. 6218510  
; GENERAL INFORMATION:  
; APPLICANT: Sharpe, Arlene H.  
; APPLICANT: Borriello, Francescopaulo  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; TITLE OF INVENTION: No. 6218510el Forms of T Cell Costimulatory Molecules  
; TITLE OF INVENTION: and Uses Therefor  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/205,697A  
; FILING DATE: 02-Mar-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: BWI-120  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400

```

; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
US-08-205-697A-39

Query Match      100.0%; Score 20; DB 3; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTCTCTGTCTAGTCCC 20
        |||
DB      57 CGTGTCTCTGTCTAGTCCC 38

RESULT 3
US-09-837-867A-39/c
; Sequence 39, Application US/09837867A
; Patent No. 6608180
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. 6608180el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: BWI-120CPADV
; CURRENT APPLICATION NUMBER: US/09/837,867A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(124)
US-09-837-867A-39

Query Match      100.0%; Score 20; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTCTCTGTCTAGTCCC 20
        |||
DB      57 CGTGTCTCTGTCTAGTCCC 38

RESULT 4
PCT-US95-02576-39/c
; Sequence 39, Application PC/TUS9502576
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules
; TITLE OF INVENTION: and Uses Therefor
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,525
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/205,697
; FILING DATE: 02-Mar-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
US-08-702-525-39

Query Match      100.0%; Score 20; DB 3; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTCTCTGTCTAGTCCC 20
        |||
DB      57 CGTGTCTCTGTCTAGTCCC 38

RESULT 2
US-08-702-525-39/c
; Sequence 39, Application US/08702525
; Patent No. 6294660
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Sharpe
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon
; APPLICANT: Nadler, Lee
; TITLE OF INVENTION: No. 6294660el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,525
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/205,697
; FILING DATE: 02-Mar-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: BWI-120CPUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
US-08-702-525-39
```

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 124 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 107..124  
PCT-US95-02576-39

Query Match 100.0%; Score 20; DB 5; Length 124;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20  
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 5  
US-09-513-999C-425/c  
; Sequence 425, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59-US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pn  
; SEQ ID NO 425  
; LENGTH: 430  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 125..430  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 125..199  
; OTHER INFORMATION: score 7.1  
; OTHER INFORMATION: seq LFVMAFLLSGAAP/LK  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 116  
; OTHER INFORMATION: r-a or g  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 121  
; OTHER INFORMATION: m-a or c  
US-09-513-999C-425

Query Match 100.0%; Score 20; DB 4; Length 430;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20  
Db 75 CGTGTGCTGTGCTAGTCCC 56

, RESULT 6

US-09-441-411-25/c  
; Sequence 25, Application US/09441411  
; Patent No. 6734172  
; GENERAL INFORMATION:  
; APPLICANT: Scholler, Nathalie B.  
; APPLICANT: Disis, Mary L.  
; APPLICANT: Hellstrom, Ingegerd  
; APPLICANT: Hellstrom, Karl Erik  
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
; FILE REFERENCE: 730033.409  
; CURRENT APPLICATION NUMBER: US/09/441,411  
; CURRENT FILING DATE: 1999-11-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 1112  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-441-411-25

Query Match 100.0%; Score 20; DB 4; Length 1112;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20  
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 7  
US-08-456-104-1/c  
; Sequence 1, Application US/08456104  
; Patent No. 5861310  
; GENERAL INFORMATION:  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; APPLICANT: Gray, Gary S.  
; TITLE OF INVENTION: TUMOR CELLS MODIFIED TO EXPRESS B7-2 AND B7-3 WITH INCREASED  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/456,104  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/101,624;  
; FILING DATE: 26-JUL-1993;  
; APPLICATION NUMBER: 08/109,393;  
; APPLICATION NUMBER: 19-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: RPI-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear



MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 107..1093  
US-08-456-104-1

Query Match 100.0%; Score 20; DB 2; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
Db 57 CGTGTCTCTGTCTAGTCCC 38

## RESULT 8

US-08-101-624-1/c  
Sequence 1, Application US/08101624  
Patent No. 5942607  
GENERAL INFORMATION:  
APPLICANT: Freeman, Gordon J.  
APPLICANT: Nadler, Lee M.  
APPLICANT: Gray, Gary S.  
TITLE OF INVENTION: No. 5942607el CTLA4/CD28 Ligands and  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/101,624  
FILING DATE: 26-JUL-1993  
PRIOR APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: RPI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1120 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 107..1093  
US-08-101-624-1

Query Match 100.0%; Score 20; DB 2; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
Db 57 CGTGTCTCTGTCTAGTCCC 38

## RESULT 9

US-08-479-744A-1/c  
Sequence 1, Application US/08479744A  
Patent No. 6084067  
GENERAL INFORMATION:  
APPLICANT: Freeman, Gordon J.  
APPLICANT: Nadler, Lee M.  
APPLICANT: Gray, Gary S.  
TITLE OF INVENTION: No. 6084067el CTLA4/CD28 Ligands and  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 60 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,744A  
FILING DATE: June 7, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/280,757  
FILING DATE: 26-JUL-1994  
APPLICATION NUMBER: 08/109,393  
FILING DATE: 28-AUG-1993  
APPLICATION NUMBER: 08/101,624  
FILING DATE: 26-JULY-1993  
APPLICATION NUMBER: 08/147,773  
FILING DATE: 3-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: RPI-004CP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1120 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 107..1093  
US-08-479-744A-1

Query Match 100.0%; Score 20; DB 3; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
Db 57 CGTGTCTCTGTCTAGTCCC 38

## RESULT 10

US-08-280-757B-1/c  
Sequence 1, Application US/08280757B  
Patent No. 6130316  
GENERAL INFORMATION:  
APPLICANT: Freeman, Gordon J.  
APPLICANT: Nadler, Lee M.  
APPLICANT: Gray, Gary S.  
APPLICANT: Greenfield, Edward

;/ TITLE OF INVENTION: No. 6130316el CTLA4/CD28 Ligands and  
;/ TITLE OF INVENTION: Uses Therefor  
;/ NUMBER OF SEQUENCES: 53  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: LAHIVE & COCKFIELD  
;/ STREET: 60 State Street, Suite 510  
;/ CITY: Boston  
;/ STATE: Massachusetts  
;/ COUNTRY: USA  
;/ ZIP: 02109  
;/  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: Patent In Release #1.0, Version #1.25  
;/  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/280,757B  
;/ FILING DATE: 26-JUL-1994  
;/ CLASSIFICATION: 435  
;/  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: 08/101,624  
;/ FILING DATE: 26-JULY-1993  
;/ APPLICATION NUMBER: 08/109,393  
;/ FILING DATE: 19-AUG-1993  
;/ APPLICATION NUMBER: 08/147,773  
;/ FILING DATE: 3-NOV-1993  
;/  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Mandragouras, Amy E.  
;/ REGISTRATION NUMBER: 36,207  
;/ REFERENCE/DOCKET NUMBER: RPI-004CP2  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (617) 227-7400  
;/ TELEFAX: (617) 227-5941  
;/  
;/ INFORMATION FOR SEQ ID NO: 1:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 1120 base pairs  
;/ TYPE: nucleic acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: CDNA  
;/ FEATURE:  
;/ NAME/KEY: CDS  
;/ LOCATION: 107..1093  
;/  
;/ US-08-280-757B-1

Query Match 100.0%; Score 20; DB 3; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20  
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 11  
US-08-205-697A-22/c  
; Sequence 22, Application US/08205697A  
; Patent No. 6218510  
; GENERAL INFORMATION:  
; APPLICANT: Sharpe, Arlene H.  
; APPLICANT: Borriello, Francescopaulo  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; TITLE OF INVENTION: No. 6218510el Forms of T Cell Costimulatory Molecules  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875

;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: ASCII Text  
;/  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/205,697A  
;/ FILING DATE: 02-Mar-1994  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Mandragouras, Amy E.  
;/ REGISTRATION NUMBER: 36,207  
;/ REFERENCE/DOCKET NUMBER: BWI-120  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (617)227-7400  
;/ TELEFAX: (617)227-5941  
;/  
;/ INFORMATION FOR SEQ ID NO: 22:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 1120 base pairs  
;/ TYPE: nucleic acid  
;/ STRANDEDNESS: double  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: CDNA  
;/ FEATURE:  
;/ NAME/KEY: CDS  
;/ LOCATION: 107..1093  
;/  
;/ US-08-205-697A-22

Query Match 100.0%; Score 20; DB 3; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20  
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 12  
US-08-702-525-22/c  
; Sequence 22, Application US/08702525  
; Patent No. 6294860  
; GENERAL INFORMATION:  
; APPLICANT: Sharpe, Sharpe  
; APPLICANT: Borriello, Francescopaulo  
; APPLICANT: Freeman, Gordon  
; APPLICANT: Nadler, Lee  
; TITLE OF INVENTION: No. 6294660el Forms of T Cell Costimulatory  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/702,525  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/205,697  
; FILING DATE: 02-Mar-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: BWI-120CPUS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 107..1093  
US-08-702-525-22

Query Match 100.0%; Score 20; DB 3; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTGTCTAGTCC 20  
Db 57 CGTGTCTGTCTAGTCC 38

RESULT 13

US-08-403-253A-3/c  
; Sequence 3, Application US/08403253A  
; Patent No. 6352694  
; GENERAL INFORMATION:

; APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.  
; APPLICANT: Gray, Gary S., Rennert, Paul D.  
; TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells  
; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/403,253A  
; FILING DATE: March 10, 1995

CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/253,964  
; FILING DATE: 3 JUNE 1994  
; APPLICATION NUMBER: US 08/073,223  
; FILING DATE: 4 JUNE 1993  
; APPLICATION NUMBER: US 08/200,947  
; FILING DATE: 23 FEB 1994  
; APPLICATION NUMBER: US 07/864,805  
; FILING DATE: 7 APR 1992  
; APPLICATION NUMBER: US 08/247,505  
; FILING DATE: 23 MAY 1994  
; APPLICATION NUMBER: US 07/864,866  
; FILING DATE: 7 APR 1992  
; APPLICATION NUMBER: US 08/218,155  
; FILING DATE: 25 MAR 1994  
; APPLICATION NUMBER: US 07/864,807  
; FILING DATE: 7 APR 1992  
; APPLICATION NUMBER: US 07/902,467  
; FILING DATE: 16 JUNE 1992  
; APPLICATION NUMBER: US 07/275,433  
; FILING DATE: 23 NOV 1988

ATTORNEY/AGENT INFORMATION:

; NAME: Mandragoras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: RPI-002CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 742-4214  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 107..1093  
US-08-403-253A-3

Query Match 100.0%; Score 20; DB 3; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTGTCTAGTCC 20  
Db 57 CGTGTCTGTCTAGTCC 38

RESULT 14

US-08-435-816A-3/c  
; Sequence 3, Application US/08435816A  
; Patent No. 6534055  
; GENERAL INFORMATION:

; APPLICANT: June, Carl H.  
; APPLICANT: Thompson, Craig B.  
; APPLICANT: Nabel, Gary J.  
; APPLICANT: Gray, Gary S.  
; APPLICANT: Rennert, Paul D.

TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells

NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,816A  
; FILING DATE: May 4, 1995

CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/403,253  
; FILING DATE: 10 MARCH 1995  
; APPLICATION NUMBER: US 08/253,964  
; FILING DATE: 3 JUNE 1994  
; APPLICATION NUMBER: US 08/073,223  
; FILING DATE: 4 JUNE 1993  
; APPLICATION NUMBER: US 08/200,947  
; FILING DATE: 23 FEB 1994  
; APPLICATION NUMBER: US 07/864,805  
; FILING DATE: 7 APR 1992  
; APPLICATION NUMBER: US 08/247,505  
; FILING DATE: 23 MAY 1994  
; APPLICATION NUMBER: US 07/864,866  
; FILING DATE: 7 APR 1992  
; APPLICATION NUMBER: US 08/218,155  
; FILING DATE: 25 MAR 1994  
; APPLICATION NUMBER: US 07/864,807  
; FILING DATE: 7 APR 1992  
; APPLICATION NUMBER: US 07/902,467  
; FILING DATE: 16 JUNE 1992  
; APPLICATION NUMBER: US 07/275,433  
; FILING DATE: 23 NOV 1988

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-002CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; US-08-435-816A-3

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 15
US-09-425-762-1/c
; Sequence 1, Application US/09425762
; Patent No. 6605279
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: No. 6605279el CTLA4/CD28 Ligands and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,762
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,744
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-004CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 107..1093
; US-09-425-762-1

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 16
US-09-837-867A-22/c
; Sequence 22, Application US/09837867A
; Patent No. 6608180
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. 8608180el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: BMI-120CPADV
; CURRENT APPLICATION NUMBER: US/09/837,867A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(1093)
; US-09-837-867A-22

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 17
US-09-206-132-1/c
; Sequence 1, Application US/09206132
; Patent No. 6723705
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: TUMOR CELLS MODIFIED TO EXPRESS B7-2 AND B7-3
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,132
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,744
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-004CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
```

/ APPLICATION NUMBER: US/09/206,132  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 08/456,104  
/ FILING DATE:  
/ APPLICATION NUMBER: 08/101,624  
/ FILING DATE: 26-JUL-1993  
/ APPLICATION NUMBER: 08/109,393  
/ APPLICATION NUMBER: 19-AUG-1993  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Mandragouras, Amy E.  
/ REGISTRATION NUMBER: 36,207  
/ REFERENCE/DOCKET NUMBER: RPI-008  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (617) 227-7400  
/ TELEFAX: (617) 227-5941  
/ INFORMATION FOR SEQ ID NO: 1:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1120 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: CDNA  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 107..1093  
US-09-206-132-1

Query Match 100.0%; Score 20; DB 4; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
DB 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 18  
PCT-US95-425-516-1/c  
; Sequence 1, Application US/09425516  
; Patent No. 6824779  
; GENERAL INFORMATION:  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; APPLICANT: Gray, Gary S.  
; TITLE OF INVENTION: No. 6824779el CTLA4/CD28 Ligands and  
; TITLE OF INVENTION: Uses Therefor  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/425,516  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/479,744  
; FILING DATE: June 7, 1995  
; APPLICATION NUMBER: 08/280,757  
; FILING DATE: 26-JUL-1994  
; APPLICATION NUMBER: 08/109,393  
; FILING DATE: 28-AUG-1993  
; APPLICATION NUMBER: 08/101,624

/ FILING DATE: 26-JULY-1993  
/ APPLICATION NUMBER: 08/147,773  
/ FILING DATE: 3-NOV-1993  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Mandragouras, Amy E.  
/ REGISTRATION NUMBER: 36,207  
/ REFERENCE/DOCKET NUMBER: RPI-004CF3  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (617) 227-7400  
/ TELEFAX: (617) 227-5941  
/ INFORMATION FOR SEQ ID NO: 1:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1120 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: CDNA  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 107..1093  
US-09-425-516-1

Query Match 100.0%; Score 20; DB 4; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20  
DB 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 19  
PCT-US95-02576-22/c  
; Sequence 22, Application PC/TUS9502576  
; GENERAL INFORMATION:  
; APPLICANT:

; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules  
; TITLE OF INVENTION: and Uses Therefor  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02576  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/205,697  
; FILING DATE: 02-Mar-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: BWI-120CPPC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS

LOCATION: 107..1093  
PCT-US95-02576-22

Query Match 100.0%; Score 20; DB 5; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20  
Db 57 CGTGTCTGTGCTAGTCC 38

## RESULT 20

US-09-949-016-12731  
; Sequence 12731, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12731  
; LENGTH: 264206  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12731

Query Match 82.0%; Score 16.4; DB 4; Length 264206;  
Best Local Similarity 94.4%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTGTGCTAGTCC 19  
Db 37623 GTGTCTGTGCTAGTCC 37640

## RESULT 21

US-09-949-016-13249  
; Sequence 13249, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13249  
; LENGTH: 264304  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13249

Query Match 82.0%; Score 16.4; DB 4; Length 264304;  
Best Local Similarity 94.4%; Pred. No. 1.7e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GTGTCTGTGCTAGTCC 19  
Db 37623 GTGTCTGTGCTAGTCC 37640

## RESULT 22

US-09-513-999C-9101  
; Sequence 9101, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 9101  
; LENGTH: 368  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 202  
; OTHER INFORMATION: k=g or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 268  
; OTHER INFORMATION: k=g or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 323  
; OTHER INFORMATION: d=a or g or t  
US-09-513-999C-9101

Query Match 80.0%; Score 16; DB 4; Length 368;  
Best Local Similarity 88.9%; Pred. No. 1.3e+02;  
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTGTGCTAGTCC 19  
Db 192 GTGTCTGTGCTAGTCC 209

## RESULT 23

US-09-799-451-340  
; Sequence 340, Application US/09799451  
; Patent No. 6783969  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Ghosh, Reena

; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6783369el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/799,451  
; CURRENT FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 948  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 340  
; LENGTH: 2304  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (61)..(1974)  
US-09-799-451-340

Query Match 79.0%; Score 15.8; DB 4; Length 2304;  
Best Local Similarity 89.5%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTGCTAGTCC 19  
|||||  
Db 651 CGTGTCTCTGTGCGAGGCC 669

RESULT 24  
US-09-949-016-17222  
; Sequence 17222, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17222  
; LENGTH: 26007  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17222

Query Match 79.0%; Score 15.8; DB 4; Length 26007;  
Best Local Similarity 89.5%; Pred. No. 2.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGTGTCTGTGCTAGTCC 20  
|||||  
Db 25893 GGGTGTCTGGGCTAGTCC 25911

RESULT 25  
US-09-949-016-13424/c  
; Sequence 13424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13424  
; LENGTH: 85850  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(85850)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13424

Query Match 79.0%; Score 15.8; DB 4; Length 85850;  
Best Local Similarity 89.5%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20  
|||||  
Db 6643 GTGTCTGTGCTATTCCC 6625

RESULT 26  
US-09-949-016-14568  
; Sequence 14568, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14568  
; LENGTH: 151295  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(151295)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14568

Query Match 79.0%; Score 15.8; DB 4; Length 151295;  
Best Local Similarity 89.5%; Pred. No. 3.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20  
|||||  
Db 143140 GTGTCTGTGCCAGTCC 143158

RESULT 27  
US-09-949-016-14569  
; Sequence 14569, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14569
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14569

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GTGTGCTGTGCTAGTCCC 20
        |||||
Db      143140 GTGTGCTGTGCTAGTCCC 143158

RESULT 28
US-09-949-016-14570
; Sequence 14570, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14570
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14570

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GTGTGCTGTGCTAGTCCC 20
        |||||
Db      143140 GTGTGCTGTGCTAGTCCC 143158

RESULT 29
US-09-949-016-14571
; Sequence 14571, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14571
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14571

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GTGTGCTGTGCTAGTCCC 20
        |||||
Db      143140 GTGTGCTGTGCTAGTCCC 143158

RESULT 30
US-09-949-016-14572
; Sequence 14572, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14572
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14572

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GTGTGCTGTGCTAGTCCC 20
        |||||
Db      143140 GTGTGCTGTGCTAGTCCC 143158

RESULT 31
US-09-949-016-14573
; Sequence 14573, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```



US-09-949-016-14573

GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14573  
; LENGTH: 393753  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(393753)  
; OTHER INFORMATION: n = A,T,C or G

Query Match 79.0%; Score 15.8; DB 4; Length 393753;  
Best Local Similarity 89.5%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20  
|||||

Db 385598 GTGTCTCTGTGCGCAGTGCC 385616

RESULT 32

US-09-949-016-14574

Sequence 14574, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14574  
; LENGTH: 393753  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(393753)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14574

Query Match 79.0%; Score 15.8; DB 4; Length 393753;  
Best Local Similarity 89.5%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20  
|||||

Db 385598 GTGTCTCTGTGCGCAGTGCC 385616

US-09-949-016-14546

Sequence 14546, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14546  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14546

Query Match 79.0%; Score 15.8; DB 4; Length 818128;  
Best Local Similarity 89.5%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20  
|||||

Db 809973 GTGTCTCTGTGCGCAGTGCC 809991

RESULT 34

US-09-949-016-14547

Sequence 14547, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14547  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14547

Query Match 79.0%; Score 15.8; DB 4; Length 818128;  
Best Local Similarity 89.5%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20  
|||||

Db 809973 GTGTCTCTGTGCGCAGTGCC 809991

```
RESULT 35
US-09-949-016-14548
; Sequence 14548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14548
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14548

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 GTGTGCTGTGCTAGTCC 20
Db  809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||

RESULT 36
US-09-949-016-14549
; Sequence 14549, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14549
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14549

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 GTGTGCTGTGCTAGTCC 20
Db  809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||
```

```
Qy  2 GTGTGCTGTGCTAGTCC 20
Db  809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||

RESULT 37
US-09-949-016-14550
; Sequence 14550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14550
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14550

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 GTGTGCTGTGCTAGTCC 20
Db  809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||

RESULT 38
US-09-949-016-14551
; Sequence 14551, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14551
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14551

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 GTGTGCTGTGCTAGTCC 20
Db  809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||
```

```
Best Local Similarity 89.5%; Pred. No. 3.4e+02; DB 4; Length 818128;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 39
US-09-949-016-14552
; Sequence 14552, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14552
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14552

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 40
US-09-949-016-14553
; Sequence 14553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14553
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14553

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 41
US-09-949-016-14554
; Sequence 14554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14554
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14554

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 42
US-09-949-016-14555
; Sequence 14555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14555
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
US-09-949-016-14553
Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 41
US-09-949-016-14554
; Sequence 14554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14554
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14554

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 42
US-09-949-016-14555
; Sequence 14555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14555
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14555

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 43
US-09-949-016-14556
; Sequence 14556, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14556
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14556

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 44
US-09-949-016-14557
; Sequence 14557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14557
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14557

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 45
US-09-949-016-14558
; Sequence 14558, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14558
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14558

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 46
US-09-949-016-14559
; Sequence 14559, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14559
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14559

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991
```

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14559  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14559

Query Match 79.0%; Score 15.8; DB 4; Length 818128;  
Best Local Similarity 89.5%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20  
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 47  
US-09-949-016-14560  
; Sequence 14560, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14560  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14560

Query Match 79.0%; Score 15.8; DB 4; Length 818128;  
Best Local Similarity 89.5%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20  
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 48  
US-09-949-016-14561  
; Sequence 14561, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14561  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14561

Query Match 79.0%; Score 15.8; DB 4; Length 818128;  
Best Local Similarity 89.5%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20  
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 49  
US-09-949-016-14562  
; Sequence 14562, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14562  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14562

Query Match 79.0%; Score 15.8; DB 4; Length 818128;  
Best Local Similarity 89.5%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20  
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 50  
US-09-949-016-14564  
; Sequence 14564, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14564
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14564

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      2  GTGTCTCTGTGCTAGTCCC 20
          |||||
Db      809973  GTGTCTCTGTGCCAGTGCC 809991
```

Search completed: March 15, 2005, 21:31:30  
Job time : 166 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 17:31:56 ; Search time 3113 Seconds

(without alignments)

244.550 Million cell updates/sec

Title: US-09-980-953-256

Perfect score: 20

Sequence: 1 cgtgtgtctgtgtagtccc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsa1:\*

9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	220	6	CD523014 AGENCOURT
C 2	20	100.0	490	6	CD698408 EST14931
C 3	20	100.0	626	5	BX371662 BX371662
C 4	20	100.0	690	5	BQ109523 imageqc.7
C 5	20	100.0	695	4	BI767024 BI767024
C 6	20	100.0	709	5	BQ109553 imageqc.7
C 7	19	95.0	132	6	CD691973 EST8496 h
C 8	19	95.0	753	4	BI906246 603063172
C 9	18.4	92.0	655	4	BI824940 603032554
C 10	18.4	92.0	995	7	CO847434 ILLUMINEX
C 11	17.4	87.0	542	1	AI992726 701493407
C 12	17.4	87.0	725	9	CC488528 CH240_320
C 13	17.4	87.0	864	8	BZ166194 CH230-463
C 14	17.4	87.0	1126	2	BF304344 CH1887245
C 15	17.4	87.0	2004	9	CL508080 SAIL 792
C 16	17	85.0	1245	5	BUI95329 AGENCOURT
C 17	16.8	84.0	363	6	BY670167 BY670167
C 18	16.8	84.0	395	9	BX963297 Reverse s
C 19	16.8	84.0	571	7	CF805725 psHA0071A
C 20	16.8	84.0	615	9	BX968677 Forward s
C 21	16.8	84.0	659	8	BZ285459 CH230-336
C 22	16.8	84.0	666	2	AW339267 xz89f02.x
C 23	16.8	84.0	696	9	CR120951 Forward s
C 24	16.8	84.0	760	9	AG353788 Mus muscu

25	16.8	84.0	840	4	BI833379
C 26	16.8	84.0	880	7	CK017551
C 27	16.8	84.0	882	6	CD754743 AGENCOURT
C 28	16.8	84.0	928	9	CR811620 GR0AAA35C
C 29	16.8	84.0	937	6	CD252254 AGENCOURT
C 30	16.8	84.0	972	9	AG426940 Mus muscu
C 31	16.8	84.0	994	4	BM048685
C 32	16.8	84.0	1010	6	CB994004 AGENCOURT
C 33	16.8	84.0	1028	5	BQ707308
C 34	16.8	84.0	1263	4	BI414634
C 35	16.8	84.0	1560	7	CF106699
C 36	16.4	82.0	178	1	AU076753
C 37	16.4	82.0	255	4	BM011282
C 38	16.4	82.0	411	7	CR546210
C 39	16.4	82.0	451	4	BG408744
C 40	16.4	82.0	481	2	AW281714
C 41	16.4	82.0	497	8	BH111553
C 42	16.4	82.0	500	6	CD589669
C 43	16.4	82.0	517	1	AA067552
C 44	16.4	82.0	544	6	CA041772
C 45	16.4	82.0	578	4	BM089797
C 46	16.4	82.0	592	6	CB504055
C 47	16.4	82.0	630	6	CA041823
C 48	16.4	82.0	636	6	CA061753
C 49	16.4	82.0	648	6	CB505188
C 50	16.4	82.0	657	1	AL847928
C 51	16.4	82.0	663	6	CA379658
C 52	16.4	82.0	684	4	BJ594673
C 53	16.4	82.0	711	4	BJ590911
C 54	16.4	82.0	730	1	AJ425708
C 55	16.4	82.0	732	4	BJ587908
C 56	16.4	82.0	756	4	BJ579066
C 57	16.4	82.0	756	4	BJ610428
C 58	16.4	82.0	784	5	BP142881
C 59	16.4	82.0	787	6	CA362363
C 60	16.4	82.0	787	9	AG402582
C 61	16.4	82.0	791	9	CNS0376M
C 62	16.4	82.0	798	9	CR080093
C 63	16.4	82.0	808	5	BP170401
C 64	16.4	82.0	817	9	CNS038AK
C 65	16.4	82.0	833	7	CNI63617
C 66	16.4	82.0	843	9	CNS01VIB
C 67	16.4	82.0	870	9	CNS031S7
C 68	16.4	82.0	1007	2	BF537884
C 69	16.4	82.0	1185	4	BG293564
C 70	16.4	82.0	1223	9	AG574348
C 71	16.4	82.0	1366	9	AG321235
C 72	16.4	82.0	1686	4	BG484292
C 73	16.4	82.0	1691	2	BF342971
C 74	16	80.0	267	8	AJ296490
C 75	16	80.0	422	9	CR224707
C 76	16	80.0	445	8	AJ2625316
C 77	16	80.0	450	8	AG680838
C 78	16	80.0	612	8	AJ427072
C 79	16	80.0	829	2	BE569899
C 80	16	80.0	859	2	BF532181
C 81	16	80.0	903	2	BE961035
C 82	16	80.0	1420	2	BF128220
C 83	16	80.0	1461	4	BG844488
C 84	16	80.0	1974	9	AG187377
C 85	15.8	79.0	139	4	BG299044
C 86	15.8	79.0	172	8	AG464084
C 87	15.8	79.0	173	6	BY566602
C 88	15.8	79.0	176	2	AW292530
C 89	15.8	79.0	188	1	AI206251
C 90	15.8	79.0	206	8	AQ639747
C 91	15.8	79.0	216	2	BE595788
C 92	15.8	79.0	225	8	AQ949378
C 93	15.8	79.0	236	2	BF221719
C 94	15.8	79.0	244	5	BQ042798
C 95	15.8	79.0	257	4	BM318712
C 96	15.8	79.0	273	1	AW611476
C 97	15.8	79.0	274	1	AI919398

BI833379	603088013
CK017551	AGENCOURT
CD754743	AGENCOURT
CR811620	GR0AAA35C
CD252254	AGENCOURT
AG426940	Mus muscu
BM048685	603628283
CB994004	AGENCOURT
BQ707308	AGENCOURT
BI414634	602989764
CF106699	PISP19 Un
AU076753	AU076753
BM011282	603635446
CR546210	DKF2P459A
BG408744	gb80006.Y
AW281714	fj53h11.x
BH111553	RPCI-24-2
CD589669	RK053A3FO
AA067552	26183 Lam
CA041772	ssalplnb5
BM089797	503647 MA
CB504055	ssalplnb5
CA041823	ssalplnb5
CA061753	ssalxgbs1
CB505188	ssalngeso
AL847928	AL847928
CA379658	658790 NC
BJ594673	BJ594673
BJ590911	BJ590911
AJ425708	AJ425708
BJ587908	BJ587908
BJ579066	BJ579066
BJ610428	BJ610428
BP142881	BP142881
CA362363	636481 NC
AG402582	Mus muscu
CR080093	Reverse s
BP170401	BP170401
AL232373	Tetraodon
CNI63617	953145 MA
AL169744	Tetraodon
AL245968	Tetraodon
BF537884	602049323
BG293564	602390214
AG574348	Mus muscu
AG321235	Mus muscu
BG484292	102402180
BF342971	602017076
AJ296490	RPCI-23-1
CR224707	Forward s
AJ2625316	1M0464A16
AG680838	HS 5488.A
AJ427072	1M0208P13
BE569899	601332108
BF532181	602073062
BE961035	601648406
BF128220	601810451
BG844488	102400680
AG187377	Fan trogl
BG299044	602397714
AQ464084	RPCI93-Dp
BY566602	BY566602
AW292530	UI-H-B12-
AI206251	QZ2F03.x
AQ639747	Z27P1-19G
BE595788	Pil 54_A0
AQ949378	Sheared D
BF221719	7141g02.x
BQ042798	UI-W-EMO-
BM318712	Pil 16_G0
AW611476	hg91a06.x
AI919398	tm83c12.x

98	15.8	79.0	280	8	AQ661130	Sheared D	c 171	15.8	79.0	527	9	CNS03JFD	AL246802 Tetradon
99	15.8	79.0	295	9	CE037312	tigr-gss-	172	15.8	79.0	528	9	TA939G11P	AL498746 T. brucei
100	15.8	79.0	297	5	BQ291932	PM3-AN009	c 173	15.8	79.0	529	8	AQ940613	Sheared D
101	15.8	79.0	307	5	BY465526	BY465526	c 174	15.8	79.0	530	8	AQ884727	HS_5517_A
102	15.8	79.0	320	2	BB586557	BB586557	175	15.8	79.0	530	8	AQ661125	Sheared D
103	15.8	79.0	329	8	AZ592777	IM0403N21	176	15.8	79.0	533	8	AQ661125	T. brucei
104	15.8	79.0	331	2	AW139807	UI-H-B11-	c 177	15.8	79.0	541	8	AQ784412	HS_3141_A
105	15.8	79.0	335	1	AA464973	zx80g11-g	c 178	15.8	79.0	542	8	AQ650723	Sheared D
106	15.8	79.0	350	2	AW409211	fb2_C2_Fe	c 179	15.8	79.0	548	8	BZ859355	CH240_232
107	15.8	79.0	352	1	AJ397018	AJ397018	c 180	15.8	79.0	550	5	BX864609	BX864609
108	15.8	79.0	355	1	AL657936	AL657936	c 181	15.8	79.0	551	8	AQ948012	Sheared D
109	15.8	79.0	361	2	BE682601	180759_MA	c 182	15.8	79.0	553	8	AQ618520	HS_5170_B
110	15.8	79.0	367	9	TA108A02P	T. brucei	c 183	15.8	79.0	554	4	BI398572	MI-P-AV1-
111	15.8	79.0	372	8	AQ640322	927P1-2H1	184	15.8	79.0	554	7	CO067284	Mdfw2064m
112	15.8	79.0	374	8	BZ882695	CH240_241	185	15.8	79.0	554	7	TA272E10Q	T. brucei
113	15.8	79.0	375	8	AZ399414	IM0165E17	186	15.8	79.0	555	8	AQ944550	Sheared D
114	15.8	79.0	379	9	TA153H09P	AL467278 T. brucei	c 187	15.8	79.0	556	9	TA393G11Q	T. brucei
115	15.8	79.0	399	1	AI347122	qp55d05.x	c 188	15.8	79.0	561	4	BI210237	EST528277
116	15.8	79.0	402	8	AQ085410	HS_2164_B	c 189	15.8	79.0	563	2	BE857086	7923h07.x
117	15.8	79.0	402	8	AQ684123	AQ085410 HS_2164_B	c 190	15.8	79.0	563	7	CN544769	UI-R-DY1-
118	15.8	79.0	407	9	CB622473	ti9r-gss-	c 191	15.8	79.0	564	2	BF078301	T. brucei
119	15.8	79.0	411	1	AI992007	w843f12.x	c 192	15.8	79.0	566	9	TA153H09Q	228527_MA
120	15.8	79.0	411	2	BE367555	P11_9_C08	c 193	15.8	79.0	567	6	CAB03739	ESG0110b.
121	15.8	79.0	411	8	AQ023239	HS_2179_B	c 194	15.8	79.0	569	5	BM887972	TMT188_Hu
122	15.8	79.0	417	7	CO695732	DG11-8911	195	15.8	79.0	569	5	BI203864	EST521904
123	15.8	79.0	420	5	BY288753	BY288753	196	15.8	79.0	571	4	AZ220932	Sheared D
124	15.8	79.0	421	1	AI768063	w146c11.x	c 197	15.8	79.0	573	8	AZ220932	Sheared D
125	15.8	79.0	423	8	AQ948675	Sheared D	c 198	15.8	79.0	575	2	BF198956	248963_MA
126	15.8	79.0	426	8	AQ023204	AQ023204 HS_2177_A	199	15.8	79.0	575	6	CB497894	omykrbna5
127	15.8	79.0	428	9	TA156G10P	AL467299 T. brucei	c 200	15.8	79.0	578	4	BF997918	MR3-GNO15
128	15.8	79.0	431	1	AU088755	AU088755	c 201	15.8	79.0	579	6	CA378490	6573h07_NC
129	15.8	79.0	431	5	BY218755	BY218755	c 202	15.8	79.0	580	5	BU877392	VO33E03_P
130	15.8	79.0	431	5	AQ637907	927P1-5G1	c 203	15.8	79.0	582	3	CNS09H1A	Single re
131	15.8	79.0	440	2	BE363891	P11_10_F0	c 204	15.8	79.0	586	6	CAB04153	ESG0112b.
132	15.8	79.0	443	6	CB789266	AMGNNUC:S	c 205	15.8	79.0	588	7	CK132617	REQ03692.3
133	15.8	79.0	444	8	BH057373	RPCI-24-2	c 206	15.8	79.0	589	5	BQ580036	ACB33-Jm8
134	15.8	79.0	445	7	CN955640	3139_102-	c 207	15.8	79.0	589	8	BZ896254	NarP7_018
135	15.8	79.0	449	5	EX091000	EX091000	c 208	15.8	79.0	608	5	EX878642	WB878642
136	15.8	79.0	449	8	BZ932053	CH240_50M	c 209	15.8	79.0	608	6	CA485093	CA485093
137	15.8	79.0	453	8	BH282546	BH282546	c 210	15.8	79.0	610	9	CL388659	RPCI44_28
138	15.8	79.0	454	9	CO984662	Forward S	c 211	15.8	79.0	611	5	BQ313895	603543071
139	15.8	79.0	455	8	AQ947640	Sheared D	c 212	15.8	79.0	611	5	BQ313895	603543071
140	15.8	79.0	456	8	AZ236621	RPCI-23-5	c 213	15.8	79.0	616	9	CE698153	ti9r-gss-
141	15.8	79.0	462	1	AI262680	qk35b06.x	c 214	15.8	79.0	617	8	BH374955	AG-ND-104
142	15.8	79.0	465	6	BY556752	BY556752	c 215	15.8	79.0	617	9	TA58D12P	T. brucei
143	15.8	79.0	467	6	CB714342	AMGNNUC:N	c 216	15.8	79.0	621	9	CL706046	FHCRC-GT-
144	15.8	79.0	467	8	AQ940637	Sheared D	c 217	15.8	79.0	632	6	CA047810	ssalbrh01
145	15.8	79.0	467	8	AQ653866	Sheared D	c 218	15.8	79.0	636	2	BB626318	BB626318
146	15.8	79.0	467	9	TA58D12Q	TA58D12Q	c 219	15.8	79.0	636	8	AQ661405	Sheared D
147	15.8	79.0	468	2	BE502327	hy22a04.x	c 220	15.8	79.0	637	4	BG385464	602454347
148	15.8	79.0	473	8	AQ066534	RPCI-23-4	c 221	15.8	79.0	637	5	BQ771765	UI-H-E21-
149	15.8	79.0	479	5	BQ394802	Sheared D	c 222	15.8	79.0	638	8	AQ659749	Sheared D
150	15.8	79.0	480	9	CG931039	MBRPH94TR	c 223	15.8	79.0	639	4	BF999529	MR2-GNO15
151	15.8	79.0	480	9	CG931039	MBRPH94TR	c 224	15.8	79.0	646	9	CR813888	GROAA38D
152	15.8	79.0	483	1	AI861847	wal3f02.x	c 225	15.8	79.0	654	2	AW745415	WS1_34_E1
153	15.8	79.0	483	7	CP433978	NIR1_31_B	c 226	15.8	79.0	654	6	CA373116	647148_NC
154	15.8	79.0	484	8	AQ762564	HS_3191_B	c 227	15.8	79.0	656	1	AI758825	ty24h06.x
155	15.8	79.0	490	8	AQ944554	Sheared D	c 228	15.8	79.0	656	5	BQ292237	ti9r-gss-
156	15.8	79.0	492	8	BH035715	RPCI-24-3	c 229	15.8	79.0	656	9	CE159452	CBI59452
157	15.8	79.0	493	8	AQ947021	Sheared D	c 230	15.8	79.0	660	6	CA386599	668228_NC
158	15.8	79.0	494	1	AI698016	we19h05.x	c 231	15.8	79.0	662	8	AQ653867	Sheared D
159	15.8	79.0	496	8	AQ947123	Sheared D	c 232	15.8	79.0	668	4	BQ736737	UI-H-E21-
160	15.8	79.0	500	1	AI367097	RK66e01.x	c 233	15.8	79.0	670	5	BQ466670	UI-H-E21-
161	15.8	79.0	502	8	AQ647248	RPCI193-Dp	c 234	15.8	79.0	675	6	CA361364	635170_NC
162	15.8	79.0	506	1	AL675840	AL675840	c 235	15.8	79.0	676	8	BZ121026	CH230-366
163	15.8	79.0	510	7	CK743721	eca01-12C	c 236	15.8	79.0	676	9	CE613987	ti9r-gss-
164	15.8	79.0	512	6	CA502614	WHB4302_F	c 237	15.8	79.0	678	8	AQ162760	mgxb0016A
165	15.8	79.0	515	9	AQ776383	HS_2162_B	c 238	15.8	79.0	679	5	BQ772276	UI-H-E21-
166	15.8	79.0	515	9	CB341227	RPCI44_26	c 239	15.8	79.0	686	4	BI454025	603174926
167	15.8	79.0	521	8	AQ659748	Sheared D	c 240	15.8	79.0	694	8	BZ338342	ia94f09.9
168	15.8	79.0	522	8	BZ338341	ia94f09.b	c 241	15.8	79.0	694	7	CN131342	EC2CAA39D
169	15.8	79.0	522	8	AQ650766	Sheared D	c 242	15.8	79.0	700	6	BY750526	BY750526
170	15.8	79.0	525	9	TA91A08P	AL459570 T. brucei	c 243	15.8	79.0	701	4	BG470159	602533861



C 244	15.8	79.0	704	8	A2852768	2M0155L16	AZ852768	2M0155L16	C 317	15.8	79.0	910	2	BF337196	BF337196	602034796
C 245	15.8	79.0	709	6	CA348319	679599 NC	CA348319	679599 NC	C 318	15.8	79.0	910	4	BQ529318	BQ529318	602558429
C 246	15.8	79.0	711	9	CL494356	SAIL 593	CL494356	SAIL 593	C 319	15.8	79.0	927	2	BF137566	BF137566	601780582
C 247	15.8	79.0	718	5	BQ257197	60374470	BQ257197	60374470	C 320	15.8	79.0	928	5	BU234333	BU234333	603341663
C 248	15.8	79.0	718	7	CV131717	L3P04h04	CV131717	L3P04h04	C 321	15.8	79.0	930	3	CL504649	CL504649	SAIL 740
C 249	15.8	79.0	718	9	CL188683	104_405_1	CL188683	104_405_1	C 322	15.8	79.0	932	3	CNS09H1B	CNS09H1B	BK058155 Single re
C 250	15.8	79.0	720	5	BU458898	603367247	BU458898	603367247	C 323	15.8	79.0	932	9	CG387059	CG387059	ZMMBHC056
C 251	15.8	79.0	722	2	BF940263	7043C12.X	BF940263	7043C12.X	C 324	15.8	79.0	934	9	CNS02SZM	CNS02SZM	AL212539 Tetraodon
C 252	15.8	79.0	725	5	BU392592	603802120	BU392592	603802120	C 325	15.8	79.0	936	9	CNS04B6X	CNS04B6X	AL282786 Tetraodon
C 253	15.8	79.0	728	5	BU271430	603374454	BU271430	603374454	C 326	15.8	79.0	943	2	BF181591	BF181591	601808893
C 254	15.8	79.0	730	9	AJ810880	Bos tauru	AJ810880	Bos tauru	C 327	15.8	79.0	947	3	CNS09P6B	CNS09P6B	BK068703 Single re
C 255	15.8	79.0	735	8	BZ561711	pac82-164	BZ561711	pac82-164	C 328	15.8	79.0	956	5	BU501856	BU501856	AGENCOURT
C 256	15.8	79.0	735	9	BA165093	Danio rer	BA165093	Danio rer	C 329	15.8	79.0	957	5	BX695893	BX695893	BX695893
C 257	15.8	79.0	740	2	BE2896237	601088011	BE2896237	601088011	C 330	15.8	79.0	965	3	CNS09P6A	CNS09P6A	BK068702 Single re
C 258	15.8	79.0	740	8	AQ038247	NLI-AC7R	AQ038247	NLI-AC7R	C 331	15.8	79.0	970	2	BE312332	BE312332	601149855
C 259	15.8	79.0	744	9	AG457200	Mus muscu	AG457200	Mus muscu	C 332	15.8	79.0	971	9	CNS04CDP	CNS04CDP	AL284326 Tetraodon
C 260	15.8	79.0	752	5	BX311609	BX311609	BX311609	BX311609	C 333	15.8	79.0	976	8	BH900876	BH900876	KG04161 D
C 261	15.8	79.0	752	9	BX134416	Danio rer	BX134416	Danio rer	C 334	15.8	79.0	1005	5	BU457646	BU457646	603369748
C 262	15.8	79.0	752	9	CR141004	Forward s	CR141004	Forward s	C 335	15.8	79.0	1010	9	CL089502	CL089502	ISB1-13D4
C 263	15.8	79.0	754	5	BX867602	BX867602	BX867602	BX867602	C 336	15.8	79.0	1017	2	BE736026	BE736026	601305626
C 264	15.8	79.0	755	9	CC906090	t026m11ba	CC906090	t026m11ba	C 337	15.8	79.0	1027	9	AG484479	AG484479	Mus muscu
C 265	15.8	79.0	758	5	BU225557	603946836	BU225557	603946836	C 338	15.8	79.0	1031	2	BF578070	BF578070	602094729
C 266	15.8	79.0	759	5	BU358134	603476987	BU358134	603476987	C 339	15.8	79.0	1036	1	AL577051	AL577051	AL577051
C 267	15.8	79.0	761	1	A1419062	tfs3f01.x	A1419062	tfs3f01.x	C 340	15.8	79.0	1042	5	BQ956399	BQ956399	AGENCOURT
C 268	15.8	79.0	762	4	BI646120	603276329	BI646120	603276329	C 341	15.8	79.0	1078	7	CF219420	CF219420	AGENCOURT
C 269	15.8	79.0	763	4	BI220614	602938707	BI220614	602938707	C 342	15.8	79.0	1101	9	CNS0020K	CNS0020K	AL061901 Drosophi1
C 270	15.8	79.0	764	8	BZ141452	CH230-512	BZ141452	CH230-512	C 343	15.8	79.0	1107	6	CA468245	CA468245	AGENCOURT
C 271	15.8	79.0	766	9	CR831873	GR0AA64D	CR831873	GR0AA64D	C 344	15.8	79.0	1176	5	BU379377	BU379377	603812875
C 272	15.8	79.0	769	9	AG541487	Mus muscu	AG541487	Mus muscu	C 345	15.8	79.0	1180	2	BF208256	BF208256	601869266
C 273	15.8	79.0	773	5	BU127833	603115077	BU127833	603115077	C 346	15.8	79.0	1245	9	CL648392	CL648392	CH213-175
C 274	15.8	79.0	773	9	CC591078	CH240_391	CC591078	CH240_391	C 347	15.8	79.0	1377	2	BE901712	BE901712	601675311
C 275	15.8	79.0	776	4	BI696288	603345668	BI696288	603345668	C 348	15.8	79.0	1405	2	BE901712	BE901712	601675311
C 276	15.8	79.0	782	5	BP696732	BP696732	BP696732	BP696732	C 349	15.8	79.0	1478	2	BP035218	BP035218	601457052
C 277	15.8	79.0	784	5	BX882052	BX882052	BX882052	BX882052	C 350	15.8	79.0	1540	3	AK053303	AK053303	Mus muscu
C 278	15.8	79.0	794	5	BU234430	603411612	BU234430	603411612	C 351	15.8	79.0	1619	2	BE967010	BE967010	601660741
C 279	15.8	79.0	802	9	CR217855	Forward s	CR217855	Forward s	C 352	15.8	79.0	1645	9	AG090038	AG090038	Pan trogl
C 280	15.8	79.0	802	9	CR217855	Forward s	CR217855	Forward s	C 353	15.8	79.0	1810	2	BE964815	BE964815	601658470
C 281	15.8	79.0	803	6	CA414748	UI-H-E20-	CA414748	UI-H-E20-	C 354	15.8	79.0	4394	3	AK034303	AK034303	Mus muscu
C 282	15.8	79.0	803	9	CC502250	CH240_341	CC502250	CH240_341	C 355	15.6	78.0	1104	1	AL572578	AL572578	AL572578
C 283	15.8	79.0	808	5	BX861947	BX861947	BX861947	BX861947	C 356	15.4	77.0	177	6	CD283725	CD283725	G39168.18
C 284	15.8	79.0	810	9	CC576533	CH240_454	CC576533	CH240_454	C 357	15.4	77.0	187	4	EG737559	EG737559	f090h04.y
C 285	15.8	79.0	810	5	BU16591	603853660	BU16591	603853660	C 358	15.4	77.0	208	9	CE772664	CE772664	tigr-ges-
C 286	15.8	79.0	811	6	CA355765	627719 NC	CA355765	627719 NC	C 359	15.4	77.0	220	1	AI617408	AI617408	zehn1523.
C 287	15.8	79.0	813	4	BI663807	603288551	BI663807	603288551	C 360	15.4	77.0	289	1	AI554094	AI554094	te55a04.x
C 288	15.8	79.0	814	9	CR810483	GR0AAA33D	CR810483	GR0AAA33D	C 361	15.4	77.0	302	4	BG794074	BG794074	UTSW SM19
C 289	15.8	79.0	822	9	BX908608	Leishmani	BX908608	Leishmani	C 362	15.4	77.0	308	9	CG464406	CG464406	Ubi 192 B
C 290	15.8	79.0	823	5	BX686992	BX686992	BX686992	BX686992	C 363	15.4	77.0	313	1	AI036033	AI036033	vz66c12.r
C 291	15.8	79.0	824	7	CN141881	WOUND1_2	CN141881	WOUND1_2	C 364	15.4	77.0	323	5	BX111311	BX111311	EX111311
C 292	15.8	79.0	828	8	BZ889083	CH240_270	BZ889083	CH240_270	C 365	15.4	77.0	332	6	CD591170	CD591170	CD591170
C 293	15.8	79.0	828	9	BX908628	Leishmani	BX908628	Leishmani	C 366	15.4	77.0	342	6	CD591175	CD591175	CD591175
C 294	15.8	79.0	830	9	CR831853	GR0AAA64D	CR831853	GR0AAA64D	C 367	15.4	77.0	351	7	F14099	F14099	ATTS5021 Ve
C 295	15.8	79.0	831	7	CK791984	AGENCOURT	CK791984	AGENCOURT	C 368	15.4	77.0	356	1	AI613889	AI613889	vh86c03.y
C 296	15.8	79.0	839	4	BG330513	602430189	BG330513	602430189	C 369	15.4	77.0	391	5	BF092008	BF092008	BP092008
C 297	15.8	79.0	840	7	CO806026	AGENCOURT	CO806026	AGENCOURT	C 370	15.4	77.0	403	5	BY002135	BY002135	BP002135
C 298	15.8	79.0	844	4	BI835046	603089276	BI835046	603089276	C 371	15.4	77.0	416	6	CD603079	CD603079	RZ145A4F0
C 299	15.8	79.0	847	4	BG966477	602832987	BG966477	602832987	C 372	15.4	77.0	420	1	AJ475490	AJ475490	AD75490
C 300	15.8	79.0	848	4	BI160086	602864116	BI160086	602864116	C 373	15.4	77.0	423	1	AV809013	AV809013	AV809013
C 301	15.8	79.0	849	1	AJ394286	Medicago	AJ394286	Medicago	C 374	15.4	77.0	425	5	BP664971	BP664971	BP664971
C 302	15.8	79.0	850	9	CR312453	CR312453	CR312453	CR312453	C 375	15.4	77.0	427	1	AV816304	AV816304	AV816304
C 303	15.8	79.0	856	5	BU228764	603797117	BU228764	603797117	C 376	15.4	77.0	439	5	BP610447	BP610447	BP610447
C 304	15.8	79.0	862	7	CK866733	AGENCOURT	CK866733	AGENCOURT	C 377	15.4	77.0	444	7	CK674974	CK674974	ZF101-P00
C 305	15.8	79.0	862	7	CR283412	CR283412	CR283412	CR283412	C 378	15.4	77.0	445	7	CV350104	CV350104	MR2-OT007
C 306	15.8	79.0	865	9	BX907321	Leishmani	BX907321	Leishmani	C 379	15.4	77.0	445	8	B53592	B53592	CIT-HSP-201
C 307	15.8	79.0	870	4	BI694506	603348095	BI694506	603348095	C 380	15.4	77.0	447	6	CD604004	CD604004	RZ151A2B1
C 308	15.8	79.0	870	7	CR439904	CR439904	CR439904	CR439904	C 381	15.4	77.0	448	1	AI485739	AI485739	EST244060
C 309	15.8	79.0	876	7	CK870622	AGENCOURT	CK870622	AGENCOURT	C 382	15.4	77.0	448	2	BE354573	BE354573	EST354573
C 310	15.8	79.0	884	5	BQ950057	AGENCOURT	BQ950057	AGENCOURT	C 383	15.4	77.0	457	1	AI964327	AI964327	EST269441
C 311	15.8	79.0	894	7	CO082911	GR_Ea47L	CO082911	GR_Ea47L	C 384	15.4	77.0	462	8	BZ207416	BZ207416	CH230-437
C 312	15.8	79.0	895	5	BX690766	BX690766	BX690766	BX690766	C 385	15.4	77.0	468	8	AQ227929	AQ227929	HS_2022_B
C 313	15.8	79.0	898	5	BQ427139	CgHem_027	BQ427139	CgHem_027	C 386	15.4	77.0	480	2	BE956442	BE956442	UI-N-BH4-
C 314	15.8	79.0	898	8	BZ118645	CH230-411	BZ118645	CH230-411	C 387	15.4	77.0	481	9	CR015405	CR015405	Forward s
C 315	15.8	79.0	903	2	BF621223	HVSME000	BF621223	HVSME000	C 388	15.4	77.0	485	6	CA113442	CA113442	SCEZLB101
C 316	15.8	79.0	904	9												



```

FEATURES
source
Location/Qualifiers
1..220
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30410101"
/tissue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_191"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggcgctcgcc); Library is oligo-dT primed
and directionally cloned. PBMC - peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGGCGGCGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 220;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 61 CGTGTGCTGTGCTAGTCCC 42

RESULT 2
LOCUS CD698408 490 bp mRNA linear EST 25-JUN-2003
DEFINITION human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD698408
VERSION CD698408.1 GI:32226772
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 490)
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: YiXin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@zsuums.edu.cn.

FEATURES
source
Location/Qualifiers
1..490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 490;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||

FEATURES
source
Location/Qualifiers
1..626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1027YM22"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 100.0%; Score 20; DB 5; Length 626;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 548 CGTGTGCTGTGCTAGTCCC 567

RESULT 4
LOCUS BQ109523 690 bp mRNA linear EST 16-APR-2002
DEFINITION IMAGE:5203544 5', mRNA sequence.
ACCESSION BQ109523
VERSION BQ109523.1 GI:20159177
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 690)
AUTHORS Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
Prange,C.K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
resequencing for verification

```

JOURNAL  
COMMENT

Unpublished (2001)  
Other ESTs: BI767024  
Contact: Prange CK  
The I.M.A.G.E. Consortium  
Lawrence Livermore National Laboratory  
Livermore, CA, USA  
Email: help@image.llnl.gov  
This read has been verified (found to hit its original self in the correct orientation), as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact info@image.llnl.gov.  
Plate: LHAM11510 row: d column: 9  
Seq primer: ml3rpl  
High quality sequence stop: 690.  
Location/Qualifiers  
1. .690  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5203544"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_122"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5203544"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_122"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."  
Query Match 100.0%; Score 20; DB 5; Length 690;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 CGTGTGCTGTGCTAGTCCC 20  
|||||  
Db 64 CGTGTGCTGTGCTAGTCCC 45  
RESULT 5  
BI767024/c  
LOCUS BI767024 695 bp mRNA linear EST 25-SEP-2001  
DEFINITION 603054234F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5203544 5', mRNA sequence.  
ACCESSION BI767024.1 GI:15758602  
VERSION BI767024.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 695)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LHAM11510 row: d column: 09

FEATURES  
source

High quality sequence stop: 637.  
Location/Qualifiers  
1. .695  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5203544"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_122"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."  
Query Match 100.0%; Score 20; DB 4; Length 695;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 CGTGTGCTGTGCTAGTCCC 20  
|||||  
Db 65 CGTGTGCTGTGCTAGTCCC 46  
RESULT 6  
BI0109553/c  
LOCUS BI0109553 709 bp mRNA linear EST 16-APR-2002  
DEFINITION imagec\_7 2001/ams58bdr81.y1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5218562 5', mRNA sequence.  
ACCESSION BI0109553  
VERSION BI0109553.1 GI:20159207  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 709)  
Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and Prange, C.K.  
The I.M.A.G.E. Consortium quality control effort: clone resequencing for verification  
Unpublished (2001)  
Contact: Prange CK  
The I.M.A.G.E. Consortium  
Lawrence Livermore National Laboratory  
Livermore, CA, USA  
Email: help@image.llnl.gov  
This read has been produced as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact info@image.llnl.gov.  
Plate: LHAM11549 row: f column: 3  
Seq primer: ml3rpl  
High quality sequence stop: 709.  
Location/Qualifiers  
1. .709  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5218562"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_118"  
/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV

## ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 695;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 CGTGTGCTGTGCTAGTCCC 20  
|||||  
Db 65 CGTGTGCTGTGCTAGTCCC 46  
RESULT 6  
BI0109553/c  
LOCUS BI0109553 709 bp mRNA linear EST 16-APR-2002  
DEFINITION imagec\_7 2001/ams58bdr81.y1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5218562 5', mRNA sequence.  
ACCESSION BI0109553  
VERSION BI0109553.1 GI:20159207  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 709)  
Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and Prange, C.K.  
The I.M.A.G.E. Consortium quality control effort: clone resequencing for verification  
Unpublished (2001)  
Contact: Prange CK  
The I.M.A.G.E. Consortium  
Lawrence Livermore National Laboratory  
Livermore, CA, USA  
Email: help@image.llnl.gov  
This read has been produced as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact info@image.llnl.gov.  
Plate: LHAM11549 row: f column: 3  
Seq primer: ml3rpl  
High quality sequence stop: 709.  
Location/Qualifiers  
1. .709  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5218562"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_118"  
/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV

FEATURES  
source

High quality sequence stop: 709.  
Location/Qualifiers  
1. .709  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5218562"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_118"  
/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 709;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 20  
|||||  
Db 63 CGTGTGCTGTGCTAGTCC 44

## RESULT 7

CD691973/c  
LOCUS CD691973 132 bp mRNA linear EST 25-JUN-2003  
DEFINITION EST8496 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD691973  
VERSION CD691973.1 GI:32214208  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 132)  
Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and  
Zeng, Y.-X.  
Transcriptional Gene Expression Profile of Human Nasopharynx

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: Yixin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@gzsums.edu.cn.

## FEATURES

## source

1..132  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_libs="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

## ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 132;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 19  
|||||  
Db 99 CGTGTGCTGTGCTAGTCC 81

## RESULT 8

BI906246/c  
LOCUS BI906246 753 bp mRNA linear EST 16-OCT-2001  
DEFINITION 603063172F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5212648 5',  
mRNA sequence.  
ACCESSION BI906246  
VERSION BI906246.1 GI:16168907  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 753)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1432 row: 1 column: 14  
High quality sequence start: 27  
High quality sequence stop: 653.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 753)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11533 row: 0 column: 17  
High quality sequence stop: 719.

## FEATURES

## source

1..753  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5212648"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_118"  
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV  
(destroyed); RNA source leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH\_MGC Library."

## ORIGIN

Query Match 95.0%; Score 19; DB 4; Length 753;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 19  
|||||  
Db 56 CGTGTGCTGTGCTAGTCC 38

## RESULT 9

BI824940/c  
LOCUS BI824940 655 bp mRNA linear EST 04-OCT-2001  
DEFINITION 603032554F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5173789 5',  
mRNA sequence.  
ACCESSION BI824940  
VERSION BI824940.1 GI:15936490  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 655)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1432 row: 1 column: 14  
High quality sequence start: 27  
High quality sequence stop: 653.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

```

FEATURES
source
Location/Qualifiers
1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173789"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 92.0%; Score 18.4; DB 4; Length 655;
Best Local Similarity 95.0%; Pred. No. 3e+02; 1; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 20
|||||
Db 98 CGTGTGCTGTGCTAGTCC 79
|||||

RESULT 10
CO647434/c
LOCUS
DEFINITION
ILLUMIGEN MC0 40481 Katze WMPB2 Macaca mulatta cDNA clone
IBUW:23771 5' similar to Bases 130 to 980 highly similar to human
CD86 (Hs:27954), mRNA sequence.
ACCESSION
CO647434
VERSION
CO647434.1 GI:50568928
KEYWORDS
EST.
SOURCE
Macaca mulatta (rhesus monkey)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
REFERENCE
1 (bases 1 to 995)
Katze M.G., Thomas M., Korth M., Iadonato S.P. and Magnus C.L.
Large-scale Rhesus Macaque cDNA Sequencing
Unpublished (2003)
JOURNAL
COMMENT
Contact: C. Magnus
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagnus@illumigen.com
Sequenced on 2004.05.28. 775 Q20 bases. Library Preparation: Prof.
Michael Katze Lab at University of Washington DNA Sequencing:
Illumigen Biosciences Inc. For further information, see
http://www.macaque.org
PCR Primers
FORWARD: CCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGAATTGGTA
Insert Length: 995 Std Error: 0.00
Plate: CL000337 row: C column: 06
Seq primer: CCTCACTAAAGGGAACAAA
POLYA=yes.
Location/Qualifiers
1. .995
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBUW:23771"
/sex="male"
/tissue_type="blood"

```

```

/cell_type="PBMC"
/dev_stage="adult"
/lab_host="Electromax DH10B"
/clone_lib="Katze WMPB2"
/notes="Vector: pDONR 222; Site 1: BsrG I; Site 2: BsrG I;
Created from CloneMiner cDNA Library Construction Kit
(catalog #18249-029)"

ORIGIN
Query Match 92.0%; Score 18.4; DB 7; Length 995;
Best Local Similarity 95.0%; Pred. No. 3.1e+02; 1; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 67 CGTGTGCTGTGCTAGTCCC 48
|||||

RESULT 11
AI992726
LOCUS
DEFINITION
701493407 A. thaliana, Ohio State clone set Arabidopsis thaliana
cDNA clone 701493407, mRNA sequence.
ACCESSION
AI992726
VERSION
AI992726.1 GI:5839631
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 542)
Chen J., Moniyama M., Chan E., Mooney M., Carroon B., Gilliland D.,
Wang X., Hillman J., Guegler K., Kim C., Doyle M., Brzoska P.,
Gorgone G., Burns D., Griffin J., Mouanoutoua M., Nguyen D.,
Tan R., Rose M., Warren B., Ton B., Kastury K., Borillo C.,
Carpio T., Policky J., Suzuki G., Argentine C., Shah S.,
Nobriga A., Murry L., Turner C., Krikorian S., Elder L. and
Hanson D.
Arabidopsis thaliana Gene Expression MicroArray
Unpublished (1999)
COMMENT
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.
Location/Qualifiers
1. .542
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="701493407"
/clone_lib="A. thaliana, Ohio State clone set"
/notes="cDNA library was made from selected clones from the
Arabidopsis thaliana Ohio State clone set."

ORIGIN
Query Match 87.0%; Score 17.4; DB 1; Length 542;
Best Local Similarity 94.7%; Pred. No. 9.1e+02; 1; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
|||||
Db 94 GTGTGCTGTGCTATCCC 112
|||||

RESULT 12
CC488528/c
LOCUS
DEFINITION
CH240_320F23.T7 CHORI-240 Bos taurus genomic clone CH240_320F23,
genomic survey sequence.

```

```

ACCESSION CC488528
VERSION CC488528.1 GI:31799360
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Holt,R., Stott,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M., Chiu,R.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
Dairymple,B.P. and Tellam,R.
TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
JOURNAL Unpublished (2003)
COMMENT Other GSSs: CH240_320F23.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bccsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 320 row: F column: 23
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..725
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_320F23"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"
ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 725;
Best Local Similarity 94.7%; Pred. No. 9.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTCC 19
|||||
Db 445 CGTGTGCTGCTAGCCC 427

RESULT 13
BZ166194 864 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-463117.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-463117, genomic survey sequence.
ACCESSION BZ166194
VERSION BZ166194.1 GI:23807245
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium [LNL]
TITLE DNA Sequencing by: Incyte Genomics, Inc.
JOURNAL Clone distribution: MGC clone distribution information can be
COMMENT found through the I.M.A.G.E. Consortium [LNL] at: image.lnl.gov
Plate: LICM1002 row: j column: 12
High quality sequence stop: 638.

REFERENCE
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsagay,G., Geer,K.,
Shvartebeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,P., de Jong,P. and Fraser,C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL Unpublished (1999)
COMMENT Other GSSs: CH230-463117.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). BAC end
plate: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 463 row: I column: 17
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..864
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-463117"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"
ORIGIN
Query Match 87.0%; Score 17.4; DB 8; Length 864;
Best Local Similarity 94.7%; Pred. No. 9.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGTCTGCTAGTCCC 20
|||||
Db 820 GTGTGTCTGCTAGTCCC 838

RESULT 14
BF304344 1126 bp mRNA linear EST 21-NOV-2000
LOCUS 601887245F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4121195 5',
DEFINITION mRNA sequence.
ACCESSION BF304344
VERSION BF304344.1 GI:11251069
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ARCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium [LNL]
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium [LNL] at: image.lnl.gov
Plate: LICM1002 row: j column: 12
High quality sequence stop: 638.

```

```

FEATURES
  source
    Location/Qualifiers
      1. 1126
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4121195"
        /tissue_type="rhabdomyosarcoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_17"
        /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
        Site 2: XhoI; cDNA made by oligo-dT priming.
        Directionally cloned into EcoRI/XhoI sites using the
        following 5' adaptor: GGCACGAG(G). Size-selected >500bp
        for average insert size 1.8kb. Library constructed by
        Ling Hong in the laboratory of Gerald M. Rubin (University
        of California, Berkeley) using ZAP-cDNA synthesis kit
        (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```

  Query Match      87.0%; Score 17.4; DB 2; Length 1126;
  Best Local Similarity 94.7%; Pred. No. 9.7e+02;
  Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 GTGTGTCGTCTGCTAGTCCC 20
    ||||| ||||| ||||| |||||
Db  889 GTGTGTCGTCTGCTAGTCCC 871

RESULT 15
CL508080/c
LOCUS
DEFINITION
  CL508080      2004 bp      DNA      linear      GSS 01-APR-2004
  SAIL_792_D02_v1 SAIL Collection Arabidopsis thaliana genomic clone
  SAIL_792_D02_v1, genomic survey sequence.
ACCESSION
  CL508080
VERSION
  CL508080.1 GI:46005400
KEYWORDS
  GSS.
SOURCE
  Arabidopsis thaliana (chale cress)
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
  1 (bases 1 to 2004)
  Sessions A., Burke E., Presting, G., Aux, G., McElver, J., Patton, D.,
  Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D.,
  Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B.,
  Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S. A.
  A high-throughput Arabidopsis reverse genetics system
  Plant Cell 14 (12), 2985-2994 (2002)
  22356987
  12468722
  Contact: Sessions A
  Applied Trait Genetics
  Syngenta Biotechnology Inc.
  3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
  Email: allen.sessions@syngenta.com
  ABRC Stock Number CS835414; T-DNA left border flanking sequences of
  Syngenta Arabidopsis Insertion Library (SAIL) lines are available
  through the Arabidopsis Biological Resource Center (ABRC).
  Sequences represent a pool of amplified genomic regions and not
  single contiguous sequences.
  Class: TDNA tagged.
  Location/Qualifiers
    1. 2004
      /organism="Arabidopsis thaliana"
      /mol_type="genomic DNA"
      /ecotype="Columbia"
      /db_xref="taxon:3702"
      /clone="SAIL_792_D02_v1"
      /clone_lib="SAIL Collection"
      /note="T-DNA left border sequences were isolated using a
      modified TAIL-PCR strategy"
```

ORIGIN

```

FEATURES
  source
    Location/Qualifiers
      1. 1245
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6169502"
        /tissue_type="melanotic melanoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_72"
        /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 2 Kb. Library constructed by Life
        Technologies."
```

ORIGIN

```

  Query Match      85.0%; Score 17; DB 5; Length 1245;
  Best Local Similarity 100.0%; Pred. No. 1.5e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 CGTGTGTCGTCTGCTAGTCT 17
    ||||| ||||| ||||| |||||
Db  1209 CGTGTGTCGTCTAGT 1225

RESULT 17
BY670167
LOCUS
DEFINITION
  BY670167      363 bp      mRNA      linear      EST 16-DEC-2002
  RIKEN full-length enriched, 14.5 days embryo df/df
  Rathke's pouches Mus musculus cDNA clone K82011G18 3', mRNA
  sequence.
ACCESSION
  BY670167
VERSION
  BY670167.1 GI:27050562
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

ORIGIN

```

FEATURES
  source
    Location/Qualifiers
      1. 1245
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6169502"
        /tissue_type="melanotic melanoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_72"
        /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 2 Kb. Library constructed by Life
        Technologies."
```

ORIGIN

```

  Query Match      85.0%; Score 17; DB 5; Length 1245;
  Best Local Similarity 100.0%; Pred. No. 1.5e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 CGTGTGTCGTCTGCTAGTCT 17
    ||||| ||||| ||||| |||||
Db  1209 CGTGTGTCGTCTAGT 1225

RESULT 17
BY670167
LOCUS
DEFINITION
  BY670167      363 bp      mRNA      linear      EST 16-DEC-2002
  RIKEN full-length enriched, 14.5 days embryo df/df
  Rathke's pouches Mus musculus cDNA clone K82011G18 3', mRNA
  sequence.
ACCESSION
  BY670167
VERSION
  BY670167.1 GI:27050562
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

ORIGIN



1 (bases 1 to 363)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nikaide, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,  
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Mateuda, H.,  
 Batalov, S., Beisel, K. W., Blake, J. A., Bratt, D., Brusci, V.,  
 Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,  
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,  
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,  
 Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konggaya, A.,  
 Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,  
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
 Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G.,  
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,  
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,  
 Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K.,  
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,  
 Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,  
 Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,  
 Rogers, J., Birney, E. and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 22354683  
 12466851  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,  
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
 Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept.  
 Human Genetics University of Michigan Medical School 4301  
 MSRB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA )  
 whose assistance we gratefully acknowledge.  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 Location/Qualifiers  
 1. .363  
 /organism="Mus musculus"  
 /mol\_type="mRNA"

/db\_xref="taxon:10090"  
 /clone="K020011G18"  
 /tissue\_type="Rathke's pouches"  
 /dev\_stage="14.5 days embryo df/df"  
 /clone\_lib="RIKEN full-length enriched, 14.5 days embryo  
 df/df Rathke's pouches"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 6; Length 363;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CGTGTGCTGTGCTAGTATCCC 20  
 DB 59 CGTGTGCTGTGCTAGTATCCC 78

RESULT 18  
 BX963297  
 LOCUS  
 DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and  
 chromosome engineering clone MHPN136115, genomic survey sequence.  
 ACCESSION BX963297  
 VERSION BX963297.1 GI:49694720  
 KEYWORDS GSS; genome survey sequence; MICR.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 395)  
 Adams, D. J., Biggs, P. J., Cox, A. V., Davies, R. M., van der Weyden, L.,  
 Jonkers, J., Smith, J., Plumb, R. W., Taylor, R. G., Nishijima, I., Yu, Y.,  
 Rogers, J. and Bradley, A.  
 Direct Submision  
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. http://www.sanger.ac.uk/MICR  
 Location/Qualifiers  
 1. .395  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHPN136115"  
 /clone\_lib="MHPN"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 9; Length 395;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CGTGTGCTGTGCTAGTATCCC 20  
 DB 305 CGTGTGCTGTGCTAGTATCCC 324

RESULT 19  
 CF805725/c  
 LOCUS  
 DEFINITION pSHA0071a11r Agriculture Canada Phytoththora sojae EST project  
 Glycine max cDNA clone SHA007a11 5, mRNA sequence.  
 ACCESSION CF805725  
 VERSION CF805725.1 GI:37993979  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 571)  
 Qutob, D., Hraber, P. T., Sobral, B. W. S. and Gijzen, M.  
 Comparative analysis of expressed sequences in Phytoththora sojae  
 Plant Physiol. 123 (1), 243-254 (2000)  
 20267956

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE

```

PUBMED
COMMENT
10806241
Contact: Tyler B
Tyler lab
VBI
1880 Pratt Dr., Blacksburg, VA 24061, USA
Tel: 540-231-7318
Email: bmtyle@vt.edu
PCR Primers
FORWARD: BK reverse primer
BACKWARD: BK reverse primer
Plate: 007 row: A column: 11
Seq primer: BK reverse primer
High quality sequence stop: 571.

FEATURES
source
1..571
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Harosoy"
/db_xref="taxon:3847"
/clone="sUA007A11"
/tissue_type="Phytophthora sojae-infected hypocotyl"
/cell_line="Phytophthora sojae culture p6497"
/dev_stages="48 hr. post infection stage"
/dev_lib="Agriculture Canada Phytophthora sojae EST project"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 84.0%; Score 16.8; DB 7; Length 571;
Best Local Similarity 90.0%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 20 CGTGTGCTGTGCTGTGCTGCC 1

RESULT 20
BX968677/c
LOCUS
DEFINITION
Forward strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN410b24, genomic survey sequence.
BX968677
VERSION
BX968677.1 GI:49700100
KEYWORDS
GSS; Genome survey sequence; MICER.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 615)
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES
source
1..615
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN410b24"
/clone_lib="MHPN"

ORIGIN
Query Match 84.0%; Score 16.8; DB 9; Length 615;
Best Local Similarity 90.0%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 179 CGTGTGCTGTGCCAGTCA 160

```

```

RESULT 21
BZ285459/c
LOCUS
DEFINITION
BZ285459 659 bp DNA linear GSS 15-OCT-2002
CH230-336K14.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-336K14, genomic survey sequence.
BZ285459
VERSION
BZ285459.1 GI:24016470
KEYWORDS
GSS.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 659)
Zhao,S., Shetty,J., Shatsman,S., Tsagaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-336K14.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_end/rat/bac_end_intro.html
Plate: 336 row: K column: 14
Seq primer: SP6
Class: BAC ends.

FEATURES
Location/Qualifiers
source
1..659
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SeNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-336K14"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SeNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match 84.0%; Score 16.8; DB 8; Length 659;
Best Local Similarity 90.0%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 359 CATGTGCTGTGCTAGGCC 340

RESULT 22
AW339267
LOCUS
DEFINITION
AW339267 666 bp mRNA linear EST 31-JAN-2000
X889f02.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2871387 3',
similar to gb:M22489 BONE MORPHOGENETIC PROTEIN 2 PRECURSOR
(HUMAN); mRNA sequence.
AW339267
VERSION
AW339267.1 GI:6835926
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE 1 (bases 1 to 666)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 455.  
 Location/Qualifiers  
 1..666  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2871387"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Lu24"  
 /note="Organ: lung; Vector: pTT3D-Pac (Pharmacina) with a modified polylinker; Plasmid DNA from the normalized library NCI\_CGAP\_Lu2 was prepared, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Ronaldo."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 666;  
 Best Local Similarity 90.0%; Pred. No. 1.8e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTCTGTCTAGTCC 20

Db 574 CGTGTGCTCTGTCTAGTCA 593

## RESULT 23

CR120951  
 LOCUS Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP74c05, genomic survey sequence.  
 DEFINITION

ACCESSION CR120951.1 GI:49868401

VERSION GSS; genome survey sequence; MICEP.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

## REFERENCE

AUTHORS 1 (bases 1 to 696)  
 Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
 Rogers,J. and Bradley,A.

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. <http://www.sanger.ac.uk/MICEP>

FEATURES Location/Qualifiers

1..696

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/clone="MHP74c05"

/clone\_lib="MHP"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 696;  
 Best Local Similarity 90.0%; Pred. No. 1.8e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTCTGTCTAGTCC 20

Db 466 CGTGTGCTCTGTCTAGTCC 485

## RESULT 24

AG353788  
 LOCUS Mus musculus molossinus DNA, clone:MSMg01-151L04.T7, genomic survey  
 DEFINITION sequence.

ACCESSION AG353788

VERSION AG353788.1 GI:47927098

KEYWORDS GSS.

SOURCE Mus musculus molossinus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

## REFERENCE

AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.

TITLE BAC end Sequences of Library MSMg01

JOURNAL Unpublished

REFERENCE 2' (bases 1 to 760)

AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.

## TITLE

Direct Submission  
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail:hattori@gsc.riken.jp URL:<http://hgp.gsc.riken.go.jp/>,  
 Tel:81-45-503-9111, Fax:81-45-503-9170)

## COMMENT

Clones are derived from the mouse BAC library MSMg01. For BAC  
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).  
 Tsukuba Institute, Bio Resource Center,  
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
 Koyadai, Tsukuba, 305-0074 Japan  
 phone: 81-298-36-9189, fax: 81-298-36-9199  
 e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : T7

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

FEATURES Location/Qualifiers

source

1..760

/organism="Mus musculus molossinus"

/mol\_type="genomic DNA"

/sub\_species="molossinus"

/db\_xref="taxon:57486"

/clone="MSMg01-151L04.T7"

/sex="male"

/tissue\_type="mixture of kidney and spleen"

/clone\_lib="MSMg01 Mouse Male BAC Library"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 760;  
 Best Local Similarity 90.0%; Pred. No. 1.8e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTCTGTCTAGTCC 20

Db 460 CGTGTGCTCTGTCTAGTCC 479

## RESULT 25

BI833379  
 LOCUS 603088013F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5227083 5',  
 DEFINITION

```

mRNA sequence.
ACCESSION BI833379
VERSION BI833379.1 GI:15944929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 840)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1571 row: i column: 04
High quality sequence stop: 835.
Location/Qualifiers
1. .840
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5227083"
/lab_host="DH10B"
/clone_lib="NIH MGC 120"
/notes="Organ: pooled pancreas and spleen; Vector:
PCMV-SPT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen) Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 84.0%; Score 16.8; DB 4; Length 840;
Best Local Similarity 90.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTATCCC 20
|||||
Db 803 CGTGTGCTGTGCAAGTCTC 822

RESULT 26
CK017551/c
LOCUS CK017551 880 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT 16543712 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7044589
5', mRNA sequence.
ACCESSION CK017551
VERSION CK017551.1 GI:38543475
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 880)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1. .882
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7044589"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC_7). Library was constructed by Open Biosystems
(Huntsville, AL)."

ORIGIN
Query Match 84.0%; Score 16.8; DB 7; Length 880;
Best Local Similarity 90.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTATCCC 20
|||||
Db 100 CGTGTGCTGTGCTGTGCTCC 81

RESULT 27
CD754743/c
LOCUS CD754743 882 bp mRNA linear EST 30-JUN-2003
DEFINITION AGENCOURT 14618643 NCI_CGAP_ZEMB2 Danio rerio cDNA clone
IMAGE:6964789 5', mRNA sequence.
ACCESSION CD754743
VERSION CD754743.1 GI:32339030
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 882)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM14605 row: e column: 12
High quality sequence start: 13
High quality sequence stop: 639.
Location/Qualifiers
1. .882
/organism="Danio rerio"

FEATURES
source
1. .882
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7044589"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC_7). Library was constructed by Open Biosystems
(Huntsville, AL)."

Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM14806 row: b column: 11
High quality sequence stop: 685.
Location/Qualifiers
1. .880
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7044589"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC_7). Library was constructed by Open Biosystems
(Huntsville, AL)."

```

```

/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6964789"
/tissue_type="embryo"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_2EMB2"
/notes="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

```

## ORIGIN

```

Query Match      84.0%; Score 16.8; DB 6; Length 882;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    ||||| ||||| |||||
Db 150 CGTGTGTGTGCTAGTCC 131

```

## RESULT 28

```

CR811620/c
LOCUS
DEFINITION
GROAAA35CB12RM1 INRA BAC Bos taurus genomic clone INRA_597D12, DNA
sequence, genomic survey sequence.

```

```

ACCESSION
CR811620
VERSION
CR811620.1
KEYWORDS
GSS.
SOURCE
Bos taurus (cow)

```

## ORGANISM

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

```

```

REFERENCE
1 (bases 1 to 928)

```

```

AUTHORS
Eggen, A., Schibler, L. and Roy, A.

```

```

TITLE
Bovine BAC End Sequences from the INRA bovine BAC library

```

```

JOURNAL
Unpublished

```

```

REFERENCE
2 (bases 1 to 928)

```

```

AUTHORS
Genoscope.

```

```

TITLE
Direct Submission

```

```

JOURNAL
Submitted (20-SEP-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

```

## COMMENT

```

Contact: Andre Eggen
Department of Animal Genetics - LCBG
INRA

```

```

78350 Jouy-en-Josas, France
Tel: 33 1 34 65 24 24
Fax: 33 1 34 65 24 78

```

```

Email: eggen@jouy.inra.fr

```

```

Clones are derived from the INRA bovine BAC library

```

```

(http://locus.jouy.inra.fr/fpc/cattle_bac_map.htm). For BAC library
availability, please contact Andre Eggen (eggen@jouy.inra.fr). This
work was undertaken as part of the International Bovine BAC

```

```

Mapping Consortium (IBBMC) by INRA (Jouy-en-Josas) and Genoscope
(Evry). Plate: 597 row: D column: 12

```

```

Seq primer: M13 Reverse

```

```

Class: BAC ends.

```

```

Location/Qualifiers

```

## FEATURES

```

source

```

```

1..928

```

```

/organism="Bos taurus"

```

```

/mol_type="genomic DNA"

```

```

/strain="breed: Holstein"

```

```

/db_xref="taxon:9913"

```

```

/clone="INRA_597D12"

```

```

/sex="Male"

```

```

/cell_type="fibroblast"

```

```

/clone_lib="INRA bovine BAC"

```

```

/notes="Vector: pBeloBAC11; Site 1: HindIII; Holstein bull;
INRA Bovine BAC library (Maie) produced by Andre
Eggen-Genoscope sequence ID : GROAAA35CB12RM1"

```

## ORIGIN

```

Query Match      84.0%; Score 16.8; DB 9; Length 928;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    ||||| ||||| |||||
Db 546 CGTGTGTGTGCTAGTCC 527

```

## RESULT 29

```

CD252254/c
LOCUS
DEFINITION
AGENCOURT 14161901 NIH MGC 181 Homo sapiens cDNA clone
IMAGE:30374576 5', mRNA sequence.

```

```

ACCESSION
CD252254
VERSION
CD252254.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```

```

ORGANISM
Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
1 (bases 1 to 937)

```

```

AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.

```

```

TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)

```

```

JOURNAL
Unpublished (1999)

```

```

COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs@mail.nih.gov

```

```

Tissue procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC Clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

Plate: NDAM434 row: d column: 09
High quality sequence stop: 16
High quality sequence stop: 508.

```

```

Location/Qualifiers

```

```

1..937

```

```

/organism="Homo sapiens"

```

```

/mol_type="mRNA"

```

```

/db_xref="taxon:9606"

```

```

/clone="IMAGE:30374576"

```

```

/tissue_type="White Matter"

```

```

/dev_stage="Unknown"

```

```

/lab_host="DH10B-Ton A ( T1 and T5 phage resistance)"

```

```

/clone_lib="NIH MGC 181"

```

```

/notes="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
(destroyed); Library is oligo-dr primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."

```

```

ORIGIN

```

```

Query Match      84.0%; Score 16.8; DB 6; Length 937;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    ||||| ||||| |||||
Db 410 CGTGTCTGTGCTATGCC 391

```

## RESULT 30

```

AG426940/c
LOCUS
DEFINITION
AG426940
Mus musculus molossinus DNA, clone:MSMg01-300H07.TJ, genomic survey
sequence.

```

AG426940  
 AG426940.1 GI:48070003  
 GSS.  
 Mus musculus molossinus  
 Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
 BAC end Sequences of Library MSMG01  
 2 (bases 1 to 972)  
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
 Direct Submission  
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail: hattori@gs.c.riken.jp, URL: http://hgp.gsc.riken.go.jp/,  
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
 Clones are derived from the mouse BAC library MSMG01. For BAC  
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).  
 Tsukuba Institute of Physical and Chemical Research (RIKEN) 3-1-1  
 Koyadai, Tsukuba, 305-0074 Japan  
 phone: 81-298-36-9189, fax: 81-298-36-9199  
 e-mail: abe@rtc.riken.jp  
 PRIMERS  
 Sequencing : TJ  
 LIBRARY  
 Vector : pBACe3.6  
 R.Site 1 : EcoRI  
 R.Site 2 : EcoRI.

FEATURES  
 source  
 1..972  
 /organism="Mus musculus molossinus"  
 /mol\_type="genomic DNA"  
 /sub\_species="molossinus"  
 /db\_xref="taxon:57485"  
 /clones="MSMG01-300H07.TJ"  
 /sex="male"  
 /tissue\_type="mixture of kidney and spleen"  
 /clone\_lib="MSMG01 Mouse Male BAC Library"

Query Match 84.0%; Score 16.8; DB 9; Length 972;  
 Best Local Similarity 90.0%; Pred. No. 1.9e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATCCC 20  
 ||||| ||||| ||||| ||||| |||||  
 Db 381 CGTGTGCTGCTAGTATCCC 362

RESULT 31  
 BM048685/c  
 LOCUS  
 DEFINITION 603628283F1 NIH\_MGC\_40 Homo sapiens CDNA clone IMAGE:5456637 5',  
 mRNA sequence.  
 BM048685  
 BM048685.1 GI:16777952  
 EST.  
 Homo sapiens (human)

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgc.nci.nih.gov/  
 1 (bases 1 to 994)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: DCTD/DPF  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCMI955 row: e column: 22  
 High quality sequence stop: 319.

FEATURES  
 source

1..994  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5456637"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /notes="Organ: prostate; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 84.0%; Score 16.8; DB 4; Length 994;  
 Best Local Similarity 90.0%; Pred. No. 1.9e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATCCC 20  
 ||||| ||||| ||||| ||||| |||||  
 Db 975 CGTGTGCTGCTAGTATCCC 956

RESULT 32  
 CB994004/c

LOCUS  
 DEFINITION AGENCOURT 13617214 NIH\_MGC\_148 Homo sapiens CDNA clone  
 IMAGE:30333973 5', mRNA sequence.  
 CB994004  
 CB994004.1 GI:30288524  
 EST.

ORIGIN

Query Match 84.0%; Score 16.8; DB 4; Length 994;  
 Best Local Similarity 90.0%; Pred. No. 1.9e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATCCC 20  
 ||||| ||||| ||||| ||||| |||||  
 Db 975 CGTGTGCTGCTAGTATCCC 956

RESULT 32  
 CB994004/c  
 LOCUS  
 DEFINITION AGENCOURT 13617214 NIH\_MGC\_148 Homo sapiens CDNA clone  
 IMAGE:30333973 5', mRNA sequence.  
 CB994004  
 CB994004.1 GI:30288524  
 EST.  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgc.nci.nih.gov/  
 1 (bases 1 to 1010)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Stefan Hanson  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
 and advice from Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDAM353 row: h column: 14  
 High quality sequence stop: 246.

FEATURES  
 source

1..1010  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30333973"  
 /tissue\_type="pre-eclampsia placenta"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH\_MGC\_148"  
 /note="Organ: placenta; Vector: pBluescriptR; Site\_1:

all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 1010;  
Best Local Similarity 90.0%; Pred. No. 1.9e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTCTAGTATCC 20  
|||||  
Db 617 CGTGTCTGTCTAGTATCC 598

## RESULT 33

B0707308/c  
LOCUS B0707308 1028 bp mRNA linear EST 16-JUL-2002  
DEFINITION AGENCOURT\_8292192 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6280950  
5', mRNA sequence.

ACCESSION B0707308  
VERSION B0707308.1 GI:21846207

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1028)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. Mark Watson

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L10M2472 row: p column: 07

High quality sequence stop: 570.

Location/Qualifiers

## FEATURES

source

1..1028  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6280950"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_113"  
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 1028;  
Best Local Similarity 90.0%; Pred. No. 1.9e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTCTAGTATCC 20  
|||||  
Db 642 CGTGTATCTTGTCTAGTATCC 623

## RESULT 34

B1414634/c

LOCUS B1414634

DEFINITION

602989764F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5145855 5',

mRNA sequence.

ACCESSION B1414634

VERSION B1414634.1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1263)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M11359 row: p column: 16

High quality sequence stop: 29

High quality sequence stop: 366.

Location/Qualifiers

source

1..1263

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:5145855"

/tissue\_type="pooled lung tumors"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NCI\_CGAP\_Lu33"

/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a

modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st

strand cDNA was prepared from mRNA obtained from pooled

lung tumors with a Not I - oligo(dT) primer [5,

TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified p773 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

Qy

1 CGTGTCTGTCTAGTATCC 20

|||||

Db

858 CGTGTCTGTCTAGTATCC 839

|||||

RESULT 35

CF106699

LOCUS

DEFINITION

PSP19 undifferentiated asexual sporangia Phycophthora infestans

cDNA, mRNA sequence.

ACCESSION CF106699

VERSION CF106699.1

KEYWORDS EST.

SOURCE

Phycophthora infestans (potato late blight agent)

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

Phycophthora.

B1414634 1263 bp mRNA linear EST 14-AUG-2001  
602989764F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5145855 5',  
mRNA sequence.

ACCESSION B1414634

VERSION B1414634.1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1263)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M11359 row: p column: 16

High quality sequence stop: 29

High quality sequence stop: 366.

Location/Qualifiers

source

1..1263

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:5145855"

/tissue\_type="pooled lung tumors"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NCI\_CGAP\_Lu33"

/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a

modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st

strand cDNA was prepared from mRNA obtained from pooled

lung tumors with a Not I - oligo(dT) primer [5,

TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified p773 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

Qy

1 CGTGTCTGTCTAGTATCC 20

|||||

Db

858 CGTGTCTGTCTAGTATCC 839

|||||

RESULT 35

CF106699

LOCUS

DEFINITION

PSP19 undifferentiated asexual sporangia Phycophthora infestans

cDNA, mRNA sequence.

ACCESSION CF106699

VERSION CF106699.1

KEYWORDS EST.

SOURCE

Phycophthora infestans (potato late blight agent)

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

Phycophthora.

```

REFERENCE
AUTHORS   Kim,K.S. and Judelson,H.S.
TITLE     Sporangia-specific gene expression in the oomycete phytopathogen
JOURNAL   Phytophthora infestans
COMMENT   Eukaryot. Cell 2 (6), 1376-1385 (2003)
Contact: Judelson HS
Department of Plant Pathology
University of California
Webber Hall, Riverside, CA 92521, USA
Tel: 909 787 4199
Fax: 909 787 4294
Email: howard.judelson@ucr.edu
MRNA induced in asexual sporangia compared to hyphae.

FEATURES
source
1..1560
/organism="Phytophthora infestans"
/mol_type="mRNA"
/strain="88069"
/db_xref="taxon:4787"
/dev_stage="Asexual sporangia from hyphae grown on rye agar"
/lab_host="DH10B"
/clone_libs="Undifferentiated asexual sporangia"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Assembly of sequence reads of overlapping cDNA clones"

ORIGIN
Query Match      84.0%; Score 16.8; DB 7; Length 1560;
Best Local Similarity 90.0%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
    |||||
Db 1479 CGTGTGCTGTGCTAGTCCC 1498

RESULT 36
AU076753/c
LOCUS       AU076753              178 bp      mRNA      linear      EST 04-MAY-2000
DEFINITION AU076753 Sugano cDNA library Homo sapiens cDNA clone kaia0289
            similar to 5'-end region of Human CTLA4 counter-receptor (B7-2)
            mRNA, mRNA sequence.
ACCESSION   AU076753
VERSION     AU076753.1 GI:7439252
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H.,
            Tsunoda,T., Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A.
            and Sugano,S.
            Statistical analysis of the 5' untranslated region of human mRNA
            using 'Oligo-Capped' cDNA libraries
            Genomics 64 (3), 286-297 (2000)
JOURNAL     20221373
MEDLINE     PUBMED
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997)
            This clone was obtained from a 'full length-enriched' cDNA library
            constructed by 'oligo-capping' method. The coding region starts
            from the 50 bp upstream to the 3'-end.
            Location/Qualifiers
            1..178
            /organism="Homo sapiens"

FEATURES
source
1..178
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_libs="NIH_MGC 47"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCCAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_libs="Sugano cDNA library"

ORIGIN
Query Match      82.0%; Score 16.4; DB 1; Length 178;
Best Local Similarity 85.0%; Pred. No. 2.5e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
    |||||
Db 79 CGGNGTCTGTGCTAGTGCC 60

RESULT 37
BM011282
LOCUS       BM011282              255 bp      mRNA      linear      EST 30-OCT-2001
DEFINITION 603635446F1 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5457923 5',
            mRNA sequence.
ACCESSION   BM011282
VERSION     BM011282.1 GI:16525636
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 255)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LHCMI958 row: k column: 12
            High quality sequence stop: 229.
            Location/Qualifiers
            1..255
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="neuroblastoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_libs="NIH_MGC 47"
            /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
                    EcoRI; cDNA made by oligo-dT priming. Directionally
                    cloned into EcoRI/XhoI sites using the following 5'
                    adaptor: GGCCAG(G). Size-selected >500bp for average
                    insert size 1.8kb. Library constructed by Ling Hong in
                    the laboratory of Gerald M. Rubin (University of
                    California, Berkeley) using ZAP-cDNA synthesis kit
                    (Stratagene) and Superscript II RT (Life Technologies).
                    Note: this is a NIH_MGC Library."

ORIGIN
Query Match      82.0%; Score 16.4; DB 4; Length 255;
Best Local Similarity 89.5%; Pred. No. 2.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGTCTGTGCTAGTCCC 20
    |||||
Db 180 GTGTGTCTGTGCTAGTCCC 198

RESULT 38
CR546210/c

```



LOCUS CR546210 411 bp mRNA linear EST 07-JUL-2004  
 DEFINITION DKFZp459A2113\_r1\_459 (synonym: pcor1) Pongo pygmaeus cDNA clone  
 ACCESSION CR546210  
 VERSION CR546210.1 GI:49898344  
 KEYWORDS EST.  
 SOURCE Pongo pygmaeus (orangutan)  
 ORGANISM Pongo pygmaeus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.  
 REFERENCE 1 (bases 1 to 411)  
 AUTHORS Ansorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B.,  
 Mewes, H.W., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and  
 Wiemann, S.  
 TITLE Pongo pygmaeus mRNA (Ansorge, W., Krieger, S., Regiert, T., et al.)  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
 Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European  
 Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. This clone  
 (DKFZp459A2113) is available at the RZPD in Berlin. Please contact  
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at <http://mips.gsf.de/projects/cdna/>.

FEATURES  
 source  
 Location/Qualifiers  
 1..411  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp459A2113"  
 /dev\_stage="adult"  
 /tissue\_type="cortex"  
 /lab\_host="DH10B"  
 /clone\_lib="459 (synonym: pcor1)"  
 /note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

ORIGIN  
 Query Match 82.0%; Score 16.4; DB 7; Length 411;  
 Best Local Similarity 94.4%; Pred. No. 2.7e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 TGTGCTGCTAGTCCC 20  
 Db 177 TGTGCTTGTGCTAGTCCC 160

RESULT 39  
 EG408744/c  
 LOCUS 451 bp mRNA linear EST 13-MAR-2001  
 DEFINITION gb80d06.y1 Moss EST library PFG Physcomitrella patens cDNA clone  
 PEP SOURCE ID:PPG\_CopyA-100811\_5', mRNA sequence.  
 ACCESSION BG408744  
 VERSION BG408744.1 GI:13315137  
 KEYWORDS EST.  
 SOURCE Physcomitrella patens  
 ORGANISM Physcomitrella patens  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.  
 REFERENCE 1 (bases 1 to 451)  
 AUTHORS Quatrano, R., Bashardes, S., Cove, D., Cuming, A., Knight, C.,  
 Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T.,  
 Underwood, K., Theising, B., Allen, M., Bowers, Y., Person, B.,  
 Swaller, T., Steptoe, M., Gibbons, M., Harvey, N., Ritter, E.,  
 Jackson, Y., McCann, R., Waterston, R. and Wilson, R.  
 TITLE Leads/Wash U Moss EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Ralph Quatrano  
 Leads/Wash U Moss EST Project

Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Libraries were constructed by Dr. Stavros Bashardes as part of the  
 Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and  
 Washington Univ. in St. Louis (USA) DNA sequencing by: Washington  
 University Genome Sequencing Center For information on obtaining a  
 clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 433.  
 Location/Qualifiers  
 1..451  
 /organism="Physcomitrella patens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3218"  
 /clone="PEP SOURCE ID:PPG\_CopyA-100811"  
 /tissue\_type="gametophore: 30 day old tissue,  
 ammonium-grown"  
 /lab\_host="DH10B"  
 /clone\_lib="Moss EST library PFG"  
 /note="Vector: pAMP1; Construction of the cDNA library was  
 performed by Dr. W. Gregg Clark using a modification of  
 the cDNA synthesis protocol developed in the laboratory of  
 Dr. Michael Lovett by Dr. Yulia Korshunova (personal  
 communication). First polyA + RNA was isolated from total  
 gametophore RNA using oligo dT magnetic beads. Following  
 this, first strand cDNA synthesis was performed on the  
 bead-bound polyA + RNA, during which an oligonucleotide  
 anchor sequence was incorporated onto the 5'-ends of the  
 cDNA. PCR amplification was then used to synthesize the  
 second strand, to amplify the double stranded DNA, and to  
 incorporate dUTP containing sequences into the ends of the  
 double stranded cDNA. This DNA was size selected and  
 cloned into pAMP1 using the CloneAMP pAMP1 System (Life  
 Technologies, GibcoBRL) for cloning amplification products  
 by a non-restriction site dependant process. The cloning  
 was directional based on sequence asymmetry introduced at  
 the ends during PCR amplification. The 3' cDNA ends are  
 proximal to the NotI site of the multiple cloning site in  
 pAMP1. This annealing mixture was transformed into  
 chemically competent DH10B cells and selected for  
 ampicillin resistant growth. The resulting clones (about  
 330,000) were pooled to make the library."

ORIGIN  
 Query Match 82.0%; Score 16.4; DB 4; Length 451;  
 Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GTGTGCTGCTAGTCCC 19  
 Db 155 GTGTGCTGCTAGTCCC 138

RESULT 40  
 AW281714/c  
 LOCUS 481 bp mRNA linear EST 26-JUL-2002  
 DEFINITION fj53h11.x1 zebrafish adult brain Danio rerio cDNA 3', mRNA  
 sequence.  
 ACCESSION AW281714  
 VERSION AW281714.1 GI:6670192  
 KEYWORDS EST.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 REFERENCE 1 (bases 1 to 481)  
 AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,  
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 WASHU Zebrafish EST Project 1998  
 Unpublished (1998)  
 Other ESTs: fj53h11.y1  
 Contact: Stephen L. Johnson  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: zbrfish@wustl.edu  
 cDNA Library Preparation: John Ngai. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)

Seq primer: -40UP  
 High quality sequence stop: 440.  
 Location/Qualifiers

#### FEATURES

source

```
1. .481
  /organism="Danio rerio"
  /mol_type="mRNA"
  /db_xref="taxon:7955"
  /sex="mixed male and female"
  /tissue_type="brain"
  /dev_stage="adult"
  /lab_host="E. coli DH10B"
  /clone_lib="zebrafish adult brain"
  /note="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI;
  Original library was constructed in lambdaZIPLOX. Mass
  excision of the cDNA library was performed to yield
  pZIPLOX plasmids. Insert check was done in original
  library."
```

#### ORIGIN

```
Query Match      82.0%; Score 16.4; DB 2; Length 481;
Best Local Similarity 94.4%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      3  TGTGTCGTGCTAGTCC 20
        |||||
Db      205 TGTGTTGCTAGTCC 188
```

```
RESULT 41
LOCUS      BH111553
DEFINITION      BH111553 497 bp DNA linear GSS 19-JUL-2001
                RPCI-24-252B2.TV RPCI-24 Mus musculus genomic clone RPCI-24-252B2,
                genomic survey sequence.
ACCESSION      BH111553
VERSION        BH111553.1 GI:14946816
KEYWORDS       GSS.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 497)
Tsagave, G., Geer, K., Krol, M., Shvartsbeyn, A., Akinret, B., Levins, M.,
Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-252B2.TJ
Contact: Shaving Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
```

Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 252 row: B column: 2  
 Seq primer: T7  
 Class: BAC ends  
 Location/Qualifiers

```
1. .497
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone_lib="RPCI-24-252B2"
  /sex="Male"
  /cell_type="Spleen/Brain"
  /clone_lib="RPCI-24"
  /note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
  RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
  library was cloned in the pTARBAC1 cloning vector at the
  BamHI sites using MboI partially digested male C57BL/6J
  DNA."
```

#### ORIGIN

```
Query Match      82.0%; Score 16.4; DB 8; Length 497;
Best Local Similarity 94.4%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2  GTGTGTCGTGCTAGTCC 19
        |||||
Db      246 GTGTGTCGTGCTAGTCC 229
```

#### RESULT 42

```
CD589669/c
LOCUS      RK053A3F03.T3 Zebrafish Kidney Marrow cDNA library Danio rerio cDNA
DEFINITION      Clone RK053A3F03 5', mRNA sequence.
ACCESSION      CD589669
VERSION        CD589669.1 GI:31771021
KEYWORDS       EST.
SOURCE         Danio rerio (zebrafish)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 500)
Song, H.D., Wu, X.Y., Sun, X.J., Zhou, Y., Liu, T.X., Zhang, G.W.,
Sheng, Y., Chen, Y., Ruan, Z., Jiang, C.L., Fan, H.Y., Look, A.T.,
Zou, L.I. and Chen, Z.
Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue
Unpublished (2003)
Contact: Chen Z.
```

#### REFERENCE

AUTHORS

State Key Lab for Medical Genomics  
 Shanghai Institute of Hematology, Ruijin Hospital Affiliated to  
 Shanghai Second Medical University  
 197 Rui Jin Road II, Shanghai 200025, P. R. China  
 Tel: 86-21-64740490  
 Fax: 86-21-64743206  
 Email: zchen@stn.sh.cn

#### TITLE

JOURNAL

COMMENT

#### FEATURES

source

```
1. .500
  /organism="Danio rerio"
  /mol_type="mRNA"
  /db_xref="taxon:7955"
  /clone="RK053A3F03"
  /dev_stage="mature"
  /clone_lib="Zebrafish Kidney Marrow cDNA library"
  /note="Organ: kidney; Vector: pBS-CMV; Site_1: XhoI;
  Site_2: EcoRI; Total RNA was extracted from the kidney"
```

tissues of mature zebrafish. The poly (A) + RNA fraction was separated from total RNA by oligo (dT) cellulose chromatography. Library was initially constructed in the lambdaZAP Express vector (Stratagene) and in vivo excised into pBS-CMV vector."

## ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 500;  
Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCC 19  
|||||  
DB 25 GTGTCTCTGTCTAGTCC 8

RESULT 43  
AA067552 517 bp mRNA linear EST 06-NOV-1997  
DEFINITION 26183 Lambda-PRL2 Arabidopsis thaliana cDNA clone 78C9T7, mRNA sequence.  
ACCESSION AA067552.1 GI:1565671  
VERSION AA067552  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 517)  
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.  
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)  
95148729  
7846151  
COMMENT Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 2213tcn@bm.cl.msu.edu  
Seq primer: 17 dye primer.  
Location/Qualifiers  
1..517  
/organism="Arabidopsis thaliana".  
/mol\_type="mRNA".  
/ecotype="Columbia".  
/db\_xref="taxon:3702"  
/clone\_lib="Lambda-PRL2"  
/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dr primed cDNA."

## ORIGIN

Query Match 82.0%; Score 16.4; DB 1; Length 517;  
Best Local Similarity 89.5%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCC 20  
|||||

DB 94 GTGTCTCTGTCTATTC 112

RESULT 44  
CA041772/c

LOCUS 544 bp mRNA linear EST 04-MAR-2003  
DEFINITION ssalpinb512124 gut Salmo salar cDNA, mRNA sequence.  
ACCESSION CA041772  
VERSION CA041772.1 GI:24342692  
KEYWORDS EST.  
SOURCE Salmo salar (Atlantic salmon)  
ORGANISM Salmo salar  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
1 (bases 1 to 544)  
GRASP Consortium, Davidson,W.S., Koop,B.F. and  
http://web.uvic.ca/cbr/grasp.  
A survey of Salmo salar transcripts from high complexity cDNA libraries  
Unpublished (2002)  
Contact: Koop BF  
Centre for Biomedical Research  
University of Victoria  
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada  
Tel: 250 472 4067  
Fax: 250 472 4075  
Email: bkooop@uvic.ca

REFERENCE  
AUTHORS

TITLE A survey of Salmo salar transcripts from high complexity cDNA libraries

JOURNAL Unpublished (2002)

COMMENT Contact: Koop BF

Centre for Biomedical Research  
University of Victoria  
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada  
Tel: 250 472 4067  
Fax: 250 472 4075  
Email: bkooop@uvic.ca  
Genome Sciences Centre, BC Cancer Agency  
cDNA preparation, sequencing and bioinformatics:  
Y Butterfield, R Kirkpatrick, J Asano, N Girm, R Guin, D Lee,  
S Lee, T Olson, P Pandoh, A Prabh, D Smailus, L Spence, J Stott,  
S Taylor, G Yang, J Schein, S Jones and M Marra.  
Location/Qualifiers  
1..544  
/organism="Salmo salar"  
/mol\_type="mRNA"  
/strain="McConnell"  
/db\_xref="taxon:8030"  
/clone\_lib="gut"  
/note="Vector: pBluescriptII-SK+; Library Creator: Matthew L Rise; Atlantic salmon tissue contributors: Carlo Biagi, Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.), Simon Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery (Crofton, B.C.), Rachel Roper (University of Victoria)"

## FEATURES

source

## ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 544;  
Best Local Similarity 94.4%; Pred. No. 2.8e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCC 19  
|||||  
DB 284 GTGTCTCTGTCTAGTCC 267

RESULT 45

BM089797/c  
LOCUS 578 bp mRNA linear EST 19-NOV-2001  
DEFINITION 503647 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION BM089797  
VERSION BM089797.1 GI:17000425  
KEYWORDS EST.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
1 (bases 1 to 578)  
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,

REFERENCE

AUTHORS

**TITLE**  
 Quackenbush, J. and Keele, J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
**JOURNAL**  
 Genome Res. 11 (4), 626-630 (2001)  
**MEDLINE**  
 21180013  
**PUBMED**  
 11282978  
**COMMENT**  
 Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smithemail.marc.usda.gov  
 Single pass sequencing. Bases called and alt trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -minmatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCCGTCAGTCAGCG  
 Plate: 5 row: D column: 18  
 Seq primer: ATTTAGGTGACACTATAG.

**FEATURES**  
 source  
 1..578  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /clone\_lib="MARC 2BOV"  
 /note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
 Library made from pooled tissue from testis, thymus,  
 semitendinosus muscle, longissimus muscle, pancreas,  
 adrenal, and endometrium."

## ORIGIN

Query Match 82.0%; Score 16.4; DB 4; Length 578;  
 Best Local Similarity 94.4%; Pred.No. 2.8e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCC 19

Db 63 GTCTGTCTGTGCTAGTCC 46

**RESULT 46**  
 CB504055/c  
 LOCUS  
 DEFINITION  
 CB504055  
 ACCESSION  
 VERSION  
 CB504055.1 GI:29315281  
 KEYWORDS  
 EST.  
 SOURCE  
 Salmo salar (Atlantic salmon)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
 1 (bases 1 to 592)

GRASP Consortium, Davidson, W.S., Koop, B.F. and  
 http://web.uvic.ca/cbr/grasp.  
 A survey of Salmo salar transcripts from high complexity cDNA  
 libraries

**JOURNAL**  
**COMMENT**  
 Unpublished (2002)  
 Contact: Koop BF  
 Centre for Biomedical Research  
 University of Victoria  
 PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada  
 Tel: 250 472 4067  
 Fax: 250 472 4075  
 Email: bkoop@vic.ca

Genome Sciences Centre, BC Cancer Agency cDNA preparation,  
 sequencing and bioinformatics: Y Butterfield, R Kirkpatrick, J  
 Asano, N Girn, R Guin, D Lee, S Lee, T Olson, P Pandoh, A Prabhu,  
 D Smailus, L Spence, J Stott, S Taylor, G Yang, J Schein, S Jones and  
 M Marra.

**FEATURES**  
 source  
 1..592  
 /organism="Salmo salar"  
 /mol\_type="mRNA"  
 /strain="McConnell"  
 /db\_xref="taxon:8030"  
 /clone\_lib="gut"  
 /note="Vector: pBlueScriptIIISK+; Library Creator: Matthew  
 L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,  
 Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.); Simon  
 Jones (PBS, Nanaimo, B.C.); Seaspring Hatchery (Crofton,  
 B.C.), Rachel Roper (University of Victoria)"

## ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 592;  
 Best Local Similarity 94.4%; Pred.No. 2.8e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCC 19

Db 395 GTGTGCTGTGCTAGTCC 378

**RESULT 47**  
 CA041823  
 LOCUS  
 DEFINITION  
 seaipnb512199 gut Salmo salar cDNA, mRNA sequence.  
 ACCESSION  
 CA041823  
 VERSION  
 CA041823.1 GI:24342743  
 KEYWORDS  
 EST.  
 SOURCE  
 Salmo salar (Atlantic salmon)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
 1 (bases 1 to 630)

GRASP Consortium, Davidson, W.S., Koop, B.F. and  
 http://web.uvic.ca/cbr/grasp.  
 A survey of Salmo salar transcripts from high complexity cDNA  
 libraries

Unpublished (2002)

Contact: Koop BF  
 Centre for Biomedical Research  
 University of Victoria  
 PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada  
 Tel: 250 472 4067  
 Fax: 250 472 4075  
 Email: bkoop@uvic.ca

Genome Sciences Centre, BC Cancer Agency  
 cDNA preparation, sequencing and bioinformatics:  
 Y Butterfield, R Kirkpatrick, J Asano, N Girn, R Guin, D Lee,  
 S Lee, T Olson, P Pandoh, A Prabhu, D Smailus, L Spence, J Stott,  
 S Taylor, G Yang, J Schein, S Jones and M Marra.

POLYA=Yes.  
**FEATURES**  
 Location/Qualifiers  
 1..630  
 /organism="Salmo salar"  
 /mol\_type="mRNA"  
 /strain="McConnell"  
 /db\_xref="taxon:8030"  
 /clone\_lib="gut"

/note="Vector: pBlueScriptIIISK+; Library Creator: Matthew  
 L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,  
 Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.); Simon  
 Jones (PBS, Nanaimo, B.C.); Seaspring Hatchery (Crofton,  
 B.C.), Rachel Roper (University of Victoria)"

## ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 630;  
 Best Local Similarity 94.4%; Pred.No. 2.9e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCC 19

```

Db      517 GTGTCTCTGTGCTGTCC 534
|||||
RESULT 48
CA061753
LOCUS      636 bp      mRNA      linear      EST 04-MAR-2003
DEFINITION ssalrbs12251 mixed_tissue Salmo salar cDNA, mRNA sequence.
ACCESSION CA061753
VERSION    CA061753.1 GI:24391996
KEYWORDS   EST.
SOURCE     Salmo salar (Atlantic salmon)
ORGANISM   Salmo salar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE  1 (bases 1 to 636)
            GRASP Consortium, Davidson,W.S., Koop,B.F. and
            http://web.uvic.ca/cbr/grasp.
            A survey of Salmo salar transcripts from high complexity cDNA
            libraries
JOURNAL    Unpublished (2002)
COMMENT    Contact: Koop BF
            Centre for Biomedical Research
            University of Victoria
            PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
            Tel: 250 472 4067
            Fax: 250 472 4075
            Email: bkoop@uvic.ca
            Genome Sciences Centre, BC Cancer Agency cDNA preparation,
            sequencing and bioinformatics: Y Asano, N Girm, D Lee,
            S Taylor, G Yang, J Schein, S Jones and M Marra.
            POLYA=Yes.

FEATURES             source
    source
    1..648
    /organism="Salmo salar"
    /mol_type="mRNA"
    /strain="McConnell"
    /db_xref="taxon:8030"
    /clone_lib="gut"
    /note="Vector: pBluescriptII-SK+; Library Creator: Matthew
    L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,
    Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.), Simon
    Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery (Crofton,
    B.C.), Rachel Roper (University of Victoria)"

ORIGIN
Query Match      82.0%; Score 16.4; DB 6; Length 648;
Best Local Similarity 94.4%; Pred. No. 2.9e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GTGTCTCTGTGCTAGTCC 19
        |||||
Db      520 GTGTCTCTGTGCTGTCC 537

RESULT 50
AL847928/C
LOCUS      657 bp      mRNA      linear      EST 26-NOV-2003
DEFINITION AL847928 XGC-egg Xenopus tropicalis cDNA clone TEGG008d01 3', mRNA
            sequence.
ACCESSION  AL847928
VERSION     AL847928.2 GI:38559484
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 657)
            Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
            Sanger Xenopus tropicalis EST project 2001 (11_2003)
            Unpublished (2003)
            On Sep 15, 2002 this sequence version replaced gi:22868193.
            Contact: Taylor R
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS SEQUENCE ID: TEGG008d01.g1kT7
            Sequencing primer: T7
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dt primed from 5ug of poly A+ RNA from egg.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
            5' end and NotI at the 3' end.
            Vector: pCS107; Site 1: EcoRI; Site 2: NotI
            Host: Escherichia coli XL1-blue.
            Location/Qualifiers
            1..657
            /organism="Xenopus tropicalis"

FEATURES             source
    source
    1..657
    /organism="Xenopus tropicalis"

```

```
/mol_type="mrna"  
/db_xref="taxon:8364"  
/clone="TEgg008d01"  
/dev_stage="egg"  
/lab_host="Escherichia coli XL1-blue"  
/clone_lib="XGC-egg"  
/note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
was oligo dt primed from sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with  
EcoRI at the 5' end and NotI at the 3' end"
```

ORIGIN

```
Query Match      82.0%; Score 16.4; DB 1; Length 657;  
Best Local Similarity 94.4%; Pred. No. 2.9e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      2  GTGTGCTGTGCTAGTCC 19  
         |||||  
Db     217  GTGTGCTGTGCTAGTTC 200
```

Search completed: March 15, 2005, 21:29:05  
Job time : 3146 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 15:16:54 ; Search time 427 Seconds  
(without alignments)  
277.271 Million cell updates/sec

Title: US-09-980-953-256  
Perfect score: 20  
Sequence: 1 cgtgtgtctgtgtagtccg  
Scoring table: IDENTITY\_NUC  
Gapop-10\_0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

- Database : N\_Geneseq\_16Dec04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002as:\*
  - 7: Geneseqn2002bs:\*
  - 8: Geneseqn2003as:\*
  - 9: Geneseqn2003bs:\*
  - 10: Geneseqn2003cs:\*
  - 11: Geneseqn2003ds:\*
  - 12: Geneseqn2004as:\*
  - 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAE33173	Human B7-
2	20	100.0	20	AD27993	Human B7-
3	20	100.0	20	ADJ54435	Human B7-
C 4	20	100.0	124	AAAT01045	Human B7-
C 5	20	100.0	430	AAAC00427	Human sec
C 6	20	100.0	1120	AAQ81351	Human B 1
C 7	20	100.0	1120	AAAT49181	Human B 1
C 8	20	100.0	1120	AAV55784	Human B7-
C 9	20	100.0	1120	AAAC84049	Human B 1
C 10	20	100.0	1120	AAAB72340	Nucleotid
C 11	20	100.0	1120	AAAD27968	Human B7-
C 12	20	100.0	1120	AAAD60974	Human CD2
C 13	20	100.0	1120	AAAD69859	Human B7-
C 14	20	100.0	1120	ADJ119363	Human B7-
C 15	20	100.0	2781	ADJ54476	Human B7-
C 16	20	100.0	2781	ADP10461	Reference
C 17	20	100.0	68001	ADJ54477	Human B7-
18	17.4	87.0	3168	ADN72456	Thale cre
19	16	80.0	368	AAAC05026	Human sec
20	16	80.0	372	AAAL01062	Human rep

21	16	80.0	372	4	ABL96529	Human tes
22	16	80.0	515	8	ABX91230	Murine ge
23	16	80.0	2175	6	ABZ21938	Human CAK
C 24	16	80.0	41434	11	ACN44256	Mouse gen
C 25	15.8	79.0	439	4	ABA09377	Human DNA
C 26	15.8	79.0	468	4	AAK82836	Human imm
C 27	15.8	79.0	1345	10	AAAD51678	Human nuc
28	15.8	79.0	2304	6	ABZ11458	Human pol
29	15.8	79.0	2304	12	ADMA43976	Novel hum
30	15.8	79.0	2434	4	AAI66465	P53 tumou
31	15.8	79.0	3545	4	ABU09670	Drosophill
32	15.8	79.0	5315	10	ADSE5103	Rat gene
33	15.8	79.0	5856	4	ABU09674	Rat gene
34	15.8	79.0	82993	6	ABX09140	Mycobacte
35	15.8	79.0	110000	4	AAI99682	Continuati
36	15.8	79.0	110000	4	AAI99683	Continuati
37	15.8	79.0	214520	10	ADL13471	Osteoearth
38	15.8	79.0	347001	12	ADP43517	Human MAD
C 39	15.4	77.0	419	10	ADSE9507	Rat gene
C 40	15.4	77.0	448	12	ADJ10940	Recombina
C 41	15.4	77.0	1644	6	ABZ12515	Arabidops
C 42	15.4	77.0	1690	6	ABL41879	Nucleotid
C 43	15.4	77.0	5819	2	AAAX20533	Polynucle
C 44	15.4	77.0	21981	9	ADA02531	Mouse Sox
C 45	15.4	77.0	21981	10	ADB72269	Mouse Sox
C 46	15.4	77.0	227448	13	ABD32841	Mouse can
47	15.4	77.0	229354	6	ABQ74179	Human cmt
48	15.4	76.0	25	6	ACI35409	Human cmt
C 49	15.2	76.0	60	6	ABN39874	Human spl
C 50	15.2	76.0	425	6	ABX44836	Bovine ES
51	15.2	76.0	461	9	ACH23811	Human adu
C 52	15.2	76.0	461	9	ADH77224	Human ova
C 53	15.2	76.0	484	5	ADI70911	Human ova
C 54	15.2	76.0	484	5	ADH70911	Bovine em
C 55	15.2	76.0	575	6	ABN73071	Human ova
C 56	15.2	76.0	603	5	ADL42460	Human ova
C 57	15.2	76.0	647	5	AAE67461	DNA encod
C 58	15.2	76.0	1005	10	ADK68335	Human pri
C 59	15.2	76.0	1005	11	ADN39183	Cancer/an
C 60	15.2	76.0	1009	4	AAI59750	Human pol
C 61	15.2	76.0	1070	4	AAI57964	Human pol
C 62	15.2	76.0	1113	3	AAZ52422	HTRM clon
C 63	15.2	76.0	1538	10	ADB53498	Primary r
C 64	15.2	76.0	1538	10	ADB53498	Neuroendo
C 65	15.2	76.0	1652	4	ABL15257	Drosophill
C 66	15.2	76.0	1740	4	ADN08471	Human sec
C 67	15.2	76.0	1822	10	ADD44890	Rat gene
C 68	15.2	76.0	1822	10	ADD44894	Rat gene
C 69	15.2	76.0	3132	4	AAD08424	Human sec
C 70	15.2	76.0	3264	12	ADG76688	Phospholi
C 71	15.2	76.0	3287	4	ABL08582	Drosophill
C 72	15.2	76.0	3287	4	ABL08582	Human zin
C 73	15.2	76.0	3792	5	AAH73793	Aspergill
C 74	15.2	76.0	3828	8	ABT18955	Aspergill
C 75	15.2	76.0	3853	6	ABQ54801	Human ova
C 76	15.2	76.0	4079	8	ABT18361	Aspergill
C 77	15.2	76.0	4178	4	ABL14128	Drosophill
C 78	15.2	76.0	4302	8	ABT20775	Aspergill
C 79	15.2	76.0	4311	4	ABL15256	Drosophill
C 80	15.2	76.0	4627	8	ADT20177	Aspergill
C 81	15.2	76.0	5348	10	ADE62932	Human gen
C 82	15.2	76.0	6079	8	ABT17767	Aspergill
C 83	15.2	76.0	6089	12	ADJ95445	Mouse cDN
C 84	15.2	76.0	6089	13	ADS86867	cDNA enco
C 85	15.2	76.0	6108	4	ABL08478	Drosophill
C 86	15.2	76.0	6627	8	ABT19581	Aspergill
C 87	15.2	76.0	6808	4	ABL08488	Drosophill
C 88	15.2	76.0	16077	4	AAK86402	Human imm
C 89	15.2	76.0	17402	4	AAK86614	Human imm
C 90	15.2	76.0	17402	4	AAK86403	Human imm
C 91	15.2	76.0	24995	12	ADQ97269	Mouse can
92	15.2	76.0	30072	11	ACN44152	Mouse gen
C 93	15.2	76.0	50000	12	ADO44025	Nucleotid

94	15.2	76.0	73725	9	ADA02690	Human FLI	167	14.8	74.0	1873	4	AAS60756	Human can
95	15.2	76.0	73725	10	ADB72428	Human FLI	c 168	14.8	74.0	2999	4	ABL05137	Drosophil
96	15.2	76.0	73725	10	ADE95938	Human FLI	c 169	14.8	74.0	3433	4	AAK81255	Human imm
97	15.2	76.0	80423	13	ABD33576	Mouse can	c 170	14.8	74.0	5647	4	ABL05136	Drosophil
98	15.2	76.0	110000	12	ADH69807_2	Continuation (3 of	171	14.8	74.0	7537	4	AAK56235	Human car
99	15.2	76.0	110000	12	ADH69807_3	Continuation (4 of	172	14.8	74.0	7537	4	AAK56235	Human car
100	15.2	76.0	110000	13	ABD33293_1	Continuation (2 of	173	14.8	74.0	7537	10	AAK5643	Human imm
101	15.2	76.0	120144	13	ABD33161_1	Abd33161 Murine can	174	14.8	74.0	7537	13	ADJ08347	Human car
102	15.2	76.0	143040	11	ABD20648	Human pul	c 175	14.8	74.0	8047	4	AAK81256	Human imm
103	15.2	76.0	143067	11	ABD20815	Human pul	c 176	14.8	74.0	8047	4	AAK81256	Human imm
104	15.2	76.0	143068	11	ABD34983	Human ade	c 177	14.8	74.0	13066	4	AAK86767	Human imm
105	15.2	76.0	143068	3	AAA35150	Human ade	c 178	14.8	74.0	29411	4	AAK76613	Drosophil
106	15.2	76.0	143068	3	AAK35150	Human ade	c 179	14.8	74.0	29411	4	AAK76613	Drosophil
107	15.2	76.0	143068	3	AAK35150	Human low	180	14.8	74.0	40304	9	ADA03014	Human NCF
108	15.2	76.0	143068	3	AAK35150	Human low	181	14.8	74.0	40304	10	ADB72752	Human NCF
109	15.2	76.0	143068	6	ABL68124	Ovary can	182	14.8	74.0	40304	12	ADM74609	Human car
110	15.2	76.0	143068	6	ABL11034	Human bre	183	14.8	74.0	40304	12	ADM74609	Human car
111	15.2	76.0	143068	10	ABZ96966	Human nuc	c 184	14.8	74.0	40429	10	ADB74376	Mycobacte
112	15.2	76.0	143068	10	ABZ96966	Human nuc	185	14.8	74.0	41322	9	AAK62633	Human CD3
113	15.2	76.0	149412	3	AAA35151	Human ade	186	14.8	74.0	44456	4	ABL09962	Drosophil
114	15.2	76.0	152739	11	ABD20816	Human pul	187	14.8	74.0	44990	9	ADA02711	Mouse Vda
115	15.2	76.0	152740	3	AAK21273	Human low	188	14.8	74.0	44990	10	ADB72449	Mouse Vda
116	15.2	76.0	152740	3	AAK21273	Human low	189	14.8	74.0	44990	10	ADB72449	Mouse Vda
117	15.2	76.0	180026	12	ADQ97752	Mouse can	c 190	14.8	74.0	60935	13	ACN37224	Human cor
118	15.2	76.0	180026	12	ADQ97752	Mouse can	c 191	14.8	74.0	60935	13	ACN37224	Human cor
119	15.2	76.0	180026	12	ADQ97752	Mouse can	c 192	14.8	74.0	60935	13	ACN37224	Human cor
120	15.2	76.0	180026	12	ADQ97752	Mouse can	c 193	14.8	74.0	60935	13	ACN37224	Human cor
121	15.2	76.0	180026	12	ADQ97752	Mouse can	c 194	14.8	74.0	60935	13	ACN37224	Human cor
122	15.2	76.0	180026	12	ADQ97752	Mouse can	c 195	14.8	74.0	60935	13	ACN37224	Human cor
123	15.2	76.0	180026	12	ADQ97752	Mouse can	c 196	14.8	74.0	60935	13	ACN37224	Human cor
124	15.2	76.0	180026	12	ADQ97752	Mouse can	c 197	14.8	74.0	60935	13	ACN37224	Human cor
125	15.2	76.0	180026	12	ADQ97752	Mouse can	c 198	14.8	74.0	60935	13	ACN37224	Human cor
126	15.2	76.0	180026	12	ADQ97752	Mouse can	c 199	14.8	74.0	60935	13	ACN37224	Human cor
127	15.2	76.0	180026	12	ADQ97752	Mouse can	c 200	14.8	74.0	60935	13	ACN37224	Human cor
128	15.2	76.0	180026	12	ADQ97752	Mouse can	c 201	14.8	74.0	60935	13	ACN37224	Human cor
129	15.2	76.0	180026	12	ADQ97752	Mouse can	c 202	14.8	74.0	60935	13	ACN37224	Human cor
130	15.2	76.0	180026	12	ADQ97752	Mouse can	c 203	14.8	74.0	60935	13	ACN37224	Human cor
131	15.2	76.0	180026	12	ADQ97752	Mouse can	c 204	14.8	74.0	60935	13	ACN37224	Human cor
132	15.2	76.0	180026	12	ADQ97752	Mouse can	c 205	14.8	74.0	60935	13	ACN37224	Human cor
133	15.2	76.0	180026	12	ADQ97752	Mouse can	c 206	14.8	74.0	60935	13	ACN37224	Human cor
134	15.2	76.0	180026	12	ADQ97752	Mouse can	c 207	14.8	74.0	60935	13	ACN37224	Human cor
135	15.2	76.0	180026	12	ADQ97752	Mouse can	c 208	14.8	74.0	60935	13	ACN37224	Human cor
136	15.2	76.0	180026	12	ADQ97752	Mouse can	c 209	14.8	74.0	60935	13	ACN37224	Human cor
137	15.2	76.0	180026	12	ADQ97752	Mouse can	c 210	14.8	74.0	60935	13	ACN37224	Human cor
138	15.2	76.0	180026	12	ADQ97752	Mouse can	c 211	14.8	74.0	60935	13	ACN37224	Human cor
139	15.2	76.0	180026	12	ADQ97752	Mouse can	c 212	14.8	74.0	60935	13	ACN37224	Human cor
140	14.8	74.0	216	6	ABQ92911	Triticum	c 213	14.4	72.0	296	12	ADQ21051	Human sof
141	14.8	74.0	232	2	AAQ77572	Human gen	c 214	14.4	72.0	315	5	ABV06855	Human pro
142	14.8	74.0	232	2	AAQ77572	Human gen	c 215	14.4	72.0	315	5	ABV06855	Human pro
143	14.8	74.0	394	6	ABN66018	Human can	c 216	14.4	72.0	328	5	ABV06556	Human sof
144	14.8	74.0	408	9	ADA29682	DNA encod	c 217	14.4	72.0	404	12	ADQ18454	Human sof
145	14.8	74.0	519	9	ACH18732	Human adu	c 218	14.4	72.0	416	8	ABX62349	Arabidops
146	14.8	74.0	529	4	AAH17160	Probe #70	c 219	14.4	72.0	420	5	ABV36507	Human pro
147	14.8	74.0	529	4	AAH17160	Probe #70	c 220	14.4	72.0	445	8	ABX43581	Bovine ES
148	14.8	74.0	529	4	AAH17160	Probe #70	c 221	14.4	72.0	470	9	ACH13379	Human foe
149	14.8	74.0	529	4	AAH17160	Probe #70	c 222	14.4	72.0	472	5	ABV36800	Human pro
150	14.8	74.0	529	4	AAH17160	Probe #70	c 223	14.4	72.0	490	9	ACH13691	Human adu
151	14.8	74.0	529	4	AAH17160	Probe #70	c 224	14.4	72.0	520	5	ABV06555	Human pro
152	14.8	74.0	529	4	AAH17160	Probe #70	c 225	14.4	72.0	525	12	ACH79688	Human gen
153	14.8	74.0	529	4	AAH17160	Probe #70	c 226	14.4	72.0	545	12	ACH79688	Human gen
154	14.8	74.0	529	4	AAH17160	Probe #70	c 227	14.4	72.0	547	12	ACH79688	Human gen
155	14.8	74.0	529	4	AAH17160	Probe #70	c 228	14.4	72.0	585	3	AAK49903	Arabidops
156	14.8	74.0	529	4	AAH17160	Probe #70	c 229	14.4	72.0	611	10	ADG62612	Rat gene
157	14.8	74.0	529	4	AAH17160	Probe #70	c 230	14.4	72.0	676	5	ABX13949	Human ner
158	14.8	74.0	529	4	AAH17160	Probe #70	c 231	14.4	72.0	749	3	ABX13949	Human ner
159	14.8	74.0	529	4	AAH17160	Probe #70	c 232	14.4	72.0	760	3	AAA14071	Human SPR
160	14.8	74.0	529	4	AAH17160	Probe #70	c 233	14.4	72.0	882	5	ABV21146	Human pro
161	14.8	74.0	529	4	AAH17160	Probe #70	c 234	14.4	72.0	882	5	ABV21146	Human pro
162	14.8	74.0	529	4	AAH17160	Probe #70	c 235	14.4	72.0	985	3	AAK49911	Arabidops
163	14.8	74.0	529	4	AAH17160	Probe #70	c 236	14.4	72.0	990	3	AAK49911	Arabidops
164	14.8	74.0	529	4	AAH17160	Probe #70	c 237	14.4	72.0	1071	3	AAK49911	Arabidops
165	14.8	74.0	529	4	AAH17160	Probe #70	c 238	14.4	72.0	1076	3	AAK49911	Arabidops
166	14.8	74.0	529	4	AAH17160	Probe #70	c 239	14.4	72.0	1178	3	AAK49911	Arabidops
167	14.8	74.0	529	4	AAH17160	Probe #70	c 240	14.4	72.0	1237	3	AAK49911	Arabidops
168	14.8	74.0	529	4	AAH17160	Probe #70	c 241	14.4	72.0	1237	3	AAK49911	Arabidops
169	14.8	74.0	529	4	AAH17160	Probe #70	c 242	14.4	72.0	1237	3	AAK49911	Arabidops
170	14.8	74.0	529	4	AAH17160	Probe #70	c 243	14.4	72.0	1237	3	AAK49911	Arabidops
171	14.8	74.0	529	4	AAH17160	Probe #70	c 244	14.4	72.0	1237	3	AAK49911	Arabidops
172	14.8	74.0	529	4	AAH17160	Probe #70	c 245	14.4	72.0	1237	3	AAK49911	Arabidops
173	14.8	74.0	529	4	AAH17160	Probe #70	c 246	14.4	72.0	1237	3	AAK49911	Arabidops
174	14.8	74.0	529	4	AAH17160	Probe #70	c 247	14.4	72.0	1237	3	AAK49911	Arabidops
175	14.8	74.0	529	4	AAH17160	Probe #70	c 248	14.4	72.0	1237	3	AAK49911	Arabidops
176	14.8	74.0	529	4	AAH17160	Probe #70	c 249	14.4	72.0	1237	3	AAK49911	Arabidops
177	14.8	74.0	529	4	AAH17160	Probe #70	c 250	14.4	72.0	1237	3	AAK49911	Arabidops
178	14.8	74.0	529	4	AAH17160	Probe #70	c 251	14.4	72.0	1237	3	AAK49911	Arabidops
179	14.8	74.0	529	4	AAH17160	Probe #70	c 252	14.4	72.0	1237	3	AAK49911	Arabidops
180	14.8	74.0	529	4	AAH17160	Probe #70	c 253	14.4	72.0	1237	3	AAK49911	Arabidops
181	14.8	74.0	529	4	AAH17160	Probe #70	c 254	14.4	72.0	1237	3	AAK49911	Arabidops
182	14.8	74.0	529	4	AAH17160	Probe #70	c 255	14.4	72.0	1237	3	AAK49911	Arabidops
183	14.8	74.0	529	4	AAH17160	Probe #70	c 256	14.4	72.0	1237	3	AAK49911	Arabidops
184	14.8	74.0	529	4	AAH17160	Probe #70	c 257	14.4	72.0	1237	3	AAK49911	Arabidops
185	14.8	74.0	529	4	AAH17160	Probe #70	c 258	14.4	72.0	1237	3	AAK49911	Arabidops
186	14.8	74.0	529	4	AAH17160	Probe #70	c 259	14.4	72.0	1237	3	AAK49911	Arabidops
187	14.8	74.0	529	4	AAH17160	Probe #70	c 260	14.4	72.0	1237	3	AAK49911	Arabidops
188	14.8	74.0	529	4	AAH17160	Probe #70	c 261	14.4	72.0	1237	3	AAK49911	Arabidops
189	14.8	74.0	529	4	AAH17160	Probe #70	c 262	14.4	72.0	1237	3	AAK49911	Arabidops



240	14.4	72.0	1237	4	AD11685	Human sec	Ad11685 Human sec	c 313	14.4	72.0	25346	12	ADQ74844	AdQ74844 Loline al
241	14.4	72.0	1237	6	ABK69781	Human sec	Adk69781 Human sec	314	14.4	72.0	25806	4	ABK6766	Abk6766 Human imm
242	14.4	72.0	1237	8	ADA40473	Human sec	Ada40473 Human sec	315	14.4	72.0	29376	4	ABL08834	AbL08834 Drosophil
243	14.4	72.0	1237	8	ACC50815	Human sec	Acc50815 Human sec	c 316	14.4	72.0	32134	4	AAI199172	Aai199172 Human exc
244	14.4	72.0	1237	8	ABZ71451	Secreted	Abz71451 Secreted	c 317	14.4	72.0	32134	5	AAI163522	Aai163522 Human kid
245	14.4	72.0	1237	9	ADB91410	Human sec	Adb91410 Human sec	c 318	14.4	72.0	32192	4	AAI199173	Aai199173 Human exc
246	14.4	72.0	1237	10	ADC733922	Human sec	Adc733922 Human sec	c 319	14.4	72.0	32192	5	AAI163523	Aai163523 Human kid
247	14.4	72.0	1237	10	ADA56633	Gene enco	Ada56633 Gene enco	c 320	14.4	72.0	32763	4	AAK68779	Aak68779 Human imm
248	14.4	72.0	1251	8	ADA40097	Human sec	Ada40097 Human sec	c 321	14.4	72.0	35100	4	AAK69767	Aak69767 Human imm
249	14.4	72.0	1251	8	ACC50601	Human sec	Acc50601 Human sec	c 322	14.4	72.0	35100	4	AAK65700	Aak65700 Human imm
250	14.4	72.0	1251	8	ABZ71329	Secreted	Abz71329 Secreted	c 323	14.4	72.0	35115	4	AAK69766	Aak69766 Human imm
251	14.4	72.0	1251	9	ADB91255	Human sec	Adb91255 Human sec	c 324	14.4	72.0	35115	4	AAK65699	Aak65699 Human imm
252	14.4	72.0	1251	10	ADC73672	Human sec	Adc73672 Human sec	325	14.4	72.0	35460	12	ADQ59458	AdQ59458 Human can
253	14.4	72.0	1251	10	ADA56271	Gene enco	Ada56271 Gene enco	c 326	14.4	72.0	36785	4	AAK82208	Aak82208 Human imm
254	14.4	72.0	1340	5	ABV25507	Human pro	Abv25507 Human pro	327	14.4	72.0	70549	12	ADQ97436	AdQ97436 Mouse can
255	14.4	72.0	1340	5	ABV24187	Human pro	Abv24187 Human pro	c 328	14.4	72.0	73930	12	ADQ97577	AdQ97577 Human can
256	14.4	72.0	1412	5	ABV24402	Human pro	Abv24402 Human pro	c 329	14.4	72.0	79731	12	ADQ97640	AdQ97640 Mouse can
257	14.4	72.0	1412	5	ABV24497	Human pro	Abv24497 Human pro	c 330	14.4	72.0	95400	12	ADP08388	Adp08388 Human lam
258	14.4	72.0	1639	12	ADL12794	Human sec	Adl12794 Human sec	c 331	14.4	72.0	96276	12	ADQ17634	AdQ17634 Human sof
259	14.4	72.0	1681	4	AAH17412	Human cdn	Aah17412 Human cdn	332	14.4	72.0	96596	9	ADA02504	Ada02504 Human BAC
260	14.4	72.0	1698	6	ABK69901	Human sec	Abk69901 Human sec	333	14.4	72.0	96596	10	ADB72242	AdB72242 Human BAC
261	14.4	72.0	1740	11	ADM03148	Human cdn	Adm03148 Human cdn	334	14.4	72.0	96596	10	ADB95752	AdB95752 Human BAC
262	14.4	72.0	2030	3	AAAL4074	Human SPR	Aaal4074 Human SPR	335	14.4	72.0	110000	10	ABQ84281_2	AbQ84281_2 of
263	14.4	72.0	2043	6	ABK74362	Bacillus	Abk74362 Bacillus	336	14.4	72.0	110000	10	ABQ84281_3	AbQ84281_3 of
264	14.4	72.0	2138	3	AAAL4072	Human SPR	Aaal4072 Human SPR	c 337	14.4	72.0	168198	12	ADQ59452	AdQ59452 Human can
265	14.4	72.0	2195	3	AAAL4072	Human SPR	Aaal4072 Human SPR	c 338	14.4	72.0	173810	6	ABN85752	Abn85752 Mouse chr
266	14.4	72.0	2199	4	AAI59630	Human pol	Aai59630 Human pol	c 339	14.4	72.0	176080	12	ADL08124	AdL08124 Human gen
267	14.4	72.0	2211	5	ABAL1513	Human ner	Abal1513 Human ner	340	14.4	72.0	201143	6	ABK83568	Abk83568 Human DNA
268	14.4	72.0	2254	3	ABCT8059	Human can	Abct8059 Human can	341	14.4	72.0	208765	12	ADQ97430	AdQ97430 Mouse can
269	14.4	72.0	2330	4	AAI32646	Human can	Aai32646 Human can	342	14.4	72.0	247544	12	ADQ59419	AdQ59419 Human can
270	14.4	72.0	2390	6	ABQ76001	Human ger	Abq76001 Human ger	c 343	14.4	72.0	264965	12	ADN16203	AdN16203 Human sul
271	14.4	72.0	2393	4	AAI57844	Human pol	Aai57844 Human pol	c 344	14.4	72.0	268685	6	ABS56563	AbS56563 Human SUL
272	14.4	72.0	2483	10	ADB62098	Human cdn	Adb62098 Human cdn	345	14.4	72.0	310122	13	ABD32533	AbD32533 Mouse can
273	14.4	72.0	2485	5	ABAL17501	Human ner	Abal17501 Human ner	346	14.2	71.0	25	9	ACI07954	AcI07954 Human mic
274	14.4	72.0	2485	5	ABAL15472	Human ner	Abal15472 Human ner	c 347	14.2	71.0	33	3	AAZ96553	Aaz96553 T cell an
275	14.4	72.0	2485	5	ABAL17503	Human ner	Abal17503 Human ner	c 348	14.2	71.0	50	6	ABZ00190	AbZ00190 Human leu
276	14.4	72.0	2485	5	ABAL16616	Human ner	Abal16616 Human ner	c 349	14.2	71.0	54	2	AAZ67219	Aaz67219 Mouse CD4
277	14.4	72.0	2485	5	ABAL16617	Human ner	Abal16617 Human ner	350	14.2	71.0	95	4	AAI24619	Aai24619 Probe #14
278	14.4	72.0	2530	12	ADQ22948	Human sof	Adq22948 Human sof	351	14.2	71.0	95	4	ABA69823	AbA69823 Human foe
279	14.4	72.0	2555	3	ACAT75393	Human ORF	Ac75393 Human ORF	352	14.2	71.0	95	4	AAI49921	Aai49921 Probe #18
280	14.4	72.0	2556	4	ABL20450	Drosophil	AbL20450 Drosophil	353	14.2	71.0	95	4	ABA36691	AbA36691 Probe #15
281	14.4	72.0	2558	10	ABE99949	Bacterial	AbE99949 Bacterial	354	14.2	71.0	95	4	AAK43911	Aak43911 Human bon
282	14.4	72.0	2717	5	AAH48361	Aldehyde-	Aah48361 Aldehyde-	355	14.2	71.0	95	4	AAK18025	Aak18025 Human bra
283	14.4	72.0	2725	6	ABL67209	Thyroid c	AbL67209 Thyroid c	356	14.2	71.0	95	4	ABS43564	AbS43564 Human liv
284	14.4	72.0	2725	6	ABN95683	Gene #218	Abn95683 Gene #218	357	14.2	71.0	95	6	ABS18141	AbS18141 Human gen
285	14.4	72.0	2834	6	ABA05986	Human bra	AbA05986 Human bra	c 358	14.2	71.0	121	10	ADH92823	Adh92823 Human gen
286	14.4	72.0	2903	4	ABL17160	Drosophil	AbL17160 Drosophil	c 359	14.2	71.0	133	3	AAO4914	Aao4914 Human sec
287	14.4	72.0	3138	10	ADC32235	Human nov	Adc32235 Human nov	360	14.2	71.0	148	11	ADM13155	Adm13155 Human sec
288	14.4	72.0	3182	2	AAQ70730	TATA-bind	Aaq70730 TATA-bind	361	14.2	71.0	159	3	AAZ42384	Aaz42384 Human 5'
289	14.4	72.0	3182	2	AAI42216	Human TAR	Aai42216 Human TAR	c 362	14.2	71.0	161	2	AAQ14558	Aaq14558 HCMV AD16
290	14.4	72.0	3182	2	AAI79593	TATA-bind	Aai79593 TATA-bind	c 363	14.2	71.0	161	2	AAQ14558	Aaq14558 HCMV AD16
291	14.4	72.0	3820	12	ADQ22654	Human sof	Adq22654 Human sof	c 364	14.2	71.0	177	8	ACD05668	AcD05668 cDNA enco
292	14.4	72.0	3846	11	ACN89119	Breast ca	Acn89119 Breast ca	c 365	14.2	71.0	190	2	AAH85712	Aah85712 Human ein
293	14.4	72.0	4285	6	ABA93729	Human nuc	AbA93729 Human nuc	366	14.2	71.0	216	9	ACH45516	Ach45516 Human foe
294	14.4	72.0	4285	13	ADP55325	Human PRO	Adp55325 Human PRO	367	14.2	71.0	219	3	AACT1382	AacT1382 Human sec
295	14.4	72.0	4285	10	ABT41955	Toxicity	Abt41955 Toxicity	c 368	14.2	71.0	223	4	ABA65603	AbA65603 Human foe
296	14.4	72.0	5253	4	AAO0419	Human cel	Aao0419 Human cel	369	14.2	71.0	223	4	AAI45770	Aai45770 Probe #14
297	14.4	72.0	5253	6	ABN95157	Gene #165	Abn95157 Gene #165	370	14.2	71.0	223	4	ABA47709	AbA47709 Human bre
298	14.4	72.0	5253	12	ADQ84367	Human tum	Adq84367 Human tum	371	14.2	71.0	223	4	ABA32691	AbA32691 Probe #11
299	14.4	72.0	5549	12	ADQ22217	Human sof	Adq22217 Human sof	372	14.2	71.0	223	4	AAK39754	Aak39754 Human bon
300	14.4	72.0	6174	6	ABK72675	Human pot	Abk72675 Human pot	373	14.2	71.0	223	4	AAK14006	Aak14006 Human bra
301	14.4	72.0	8764	4	AAAL07331	Human rep	Aaal07331 Human rep	374	14.2	71.0	223	4	ABS39341	AbS39341 Human liv
302	14.4	72.0	9387	4	ABL28998	Drosophil	AbL28998 Drosophil	375	14.2	71.0	223	5	AAI06259	Aai06259 Probe #62
303	14.4	72.0	9425	4	ABL28994	Drosophil	AbL28994 Drosophil	376	14.2	71.0	223	3	ABS13848	AbS13848 Human gen
304	14.4	72.0	10445	4	AAAL03692	Human rep	Aaal03692 Human rep	c 377	14.2	71.0	238	3	AACT13151	AacT13151 Human sec
305	14.4	72.0	13909	4	ABL20392	Drosophil	AbL20392 Drosophil	c 378	14.2	71.0	240	6	ABN16239	Abn16239 Human ORF
306	14.4	72.0	14747	4	AAF63406	Human CD3	Aaf63406 Human CD3	c 379	14.2	71.0	245	8	AAI29027	Aai29027 Colon tum
307	14.4	72.0	15977	4	AAF63407	Human CD3	Aaf63407 Human CD3	c 380	14.2	71.0	245	8	ABZ33213	AbZ33213 Human col
308	14.4	72.0	17800	6	ABK72669	Human pot	Abk72669 Human pot	c 381	14.2	71.0	264	2	AAT23660	Aat23660 Human gen
309	14.4	72.0	18110	6	ABK15357	Human int	Abk15357 Human int	c 382	14.2	71.0	289	6	ABN17501	Abn17501 Human ORF
310	14.4	72.0	18340	5	AAAL19368	Mammalian	Aaal19368 Mammalian	c 383	14.2	71.0	294	6	ABN19677	Abn19677 Human ORF
311	14.4	72.0	20835	4	AAK86765	Human imm	Aak86765 Human imm	c 384	14.2	71.0	303	4	AAI82002	Aai82002 Human pol
312	14.4	72.0	21704	10	ADC86168	Human GPC	Adc86168 Human GPC	c 385	14.2	71.0	314	9	ACH31804	Ach31804 Human bon



XX WPI; 2001-049991/06.  
 XX Novel compound for diagnosing, preventing and treating immune disorders,  
 PT comprising an oligonucleotide that specifically hybridizes with a nucleic  
 PT acid sequence encoding B7 protein.  
 XX  
 PS Claim 11; Page 101; 162pp; English.  
 XX  
 CC The present invention provides sequences of antisense oligonucleotides  
 CC targeted at the murine and human B7-1 and B7-2 coding and mRNA sequences.  
 CC The antisense sequences have phosphorothioate backbones and some  
 CC nucleotides are 2'-methoxyethoxy residues. The sequences can be used in  
 CC the treatment of inflammatory and autoimmune disorders, including asthma,  
 CC juvenile diabetes mellitus, myasthenia gravis, Graves' disease,  
 CC rheumatoid arthritis, allograft rejection, inflammatory bowel disease,  
 CC multiple sclerosis, psoriasis, systemic lupus erythematosus, contact  
 CC dermatitis, rhinitis, allergies and cancer  
 XX  
 SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 4; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGTGTCTCTGCTAGTCC 20  
 |||||  
 DB 1 CGTGTCTCTGCTAGTCC 20  
 |||||  
 RESULT 2  
 ADE27993  
 ID ADE27993 standard; DNA; 20 BP.  
 XX  
 AC ADE27993;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human B7-2 targeted oligonucleotide SEQ ID 255.  
 XX  
 KW ss; human; B7-2; inflammatory skin disorder; antisense; psoriasis;  
 KW contact dermatitis; atopic dermatitis; seborrheic dermatitis;  
 KW nummular dermatitis; generalised exfoliative dermatitis; eczema;  
 KW critical costimulatory molecule.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN US2003176374-A1.  
 XX  
 PD 18-SEP-2003.  
 XX  
 PF 09-MAY-2001; 2001US-00851871.  
 XX  
 PR 31-DEC-1996; 96US-00777266.  
 PR 04-JUN-1999; 99US-00326186.  
 PR 25-MAY-2000; 2000WO-US014471.  
 XX  
 PA (BENN/) BENNETT C F.  
 PA (VICK/) VICKERS T A.  
 PA (KARR/) KARRAS J G.  
 XX  
 PI Bennett CF, Vickers TA, Karras JG;  
 XX  
 DR WPI; 2003-863863/80.  
 XX  
 XX Treating an inflammatory skin disorder such as psoriasis comprises  
 PT topically applying an antisense compound targeted to the nucleic acid  
 PT encoding human B7 protein.  
 XX  
 XX Example 20; SEQ ID NO 255; 88pp; English.  
 PS  
 CC The invention relates to a method of treating an inflammatory skin

CC disorder in an individual by topically applying an antisense compound  
 CC targeted to a nucleic acid molecule encoding a human B7 protein. The  
 CC invention is for treating an inflammatory skin disorder in individual.  
 CC The skin disorder is psoriasis, contact dermatitis, atopic dermatitis,  
 CC seborrheic dermatitis, nummular dermatitis, generalised exfoliative  
 CC dermatitis or eczema. The invention effectively modulates critical  
 CC costimulatory molecules such as the B7 protein. The present sequence  
 CC represents a human B7-2 targeted oligonucleotide.  
 XX  
 SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 10; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGTGTCTCTGCTAGTCC 20  
 |||||  
 DB 1 CGTGTCTCTGCTAGTCC 20  
 |||||  
 RESULT 3  
 ADJ54435  
 ID ADJ54435 standard; DNA; 20 BP.  
 XX  
 AC ADJ54435;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Human B7-2 DNA antisense oligonucleotide #63.  
 XX  
 KW Airway hyperresponsiveness; pulmonary inflammation;  
 KW antisense oligonucleotide; human; B7 protein; B7-2; asthma;  
 KW antiasthmatic; antiinflammatory; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004023917-A1.  
 XX  
 PD 05-FEB-2004.  
 XX  
 PF 23-MAY-2003; 2003US-00444206.  
 XX  
 PR 31-DEC-1996; 96US-00777266.  
 PR 04-JUN-1999; 99US-00326186.  
 PR 25-MAY-2000; 2000WO-US014471.  
 PR 09-MAY-2001; 2001US-00851871.  
 XX  
 PA (BENN/) BENNETT C F.  
 PA (VICK/) VICKERS T A.  
 PA (KARR/) KARRAS J G.  
 XX  
 PI Bennett CF, Vickers TA, Karras JG;  
 XX  
 DR WPI; 2004-132608/13.  
 XX  
 XX Treating airway hyperresponsiveness or pulmonary inflammation comprises  
 PT administering an antisense compound targeted to a nucleic acid molecule  
 PT encoding a human B7 protein to the individual.  
 XX  
 PS Example 20; SEQ ID NO 255; 182pp; English.  
 XX  
 XX The invention relates to a method for treating airway hyperresponsiveness  
 CC or pulmonary inflammation in an individual comprising administering an  
 CC antisense compound targeted to a nucleic acid molecule encoding a human  
 CC B7 protein. The invention also relates to a method of inhibiting  
 CC expression of a nucleic acid molecule encoding B7-1 or B7-2. The  
 CC antisense compound is an antisense oligonucleotide which has a modified  
 CC sugar moiety and nucleobase. The human B7 protein is human B7-1 or B7-2  
 CC protein or both. The compound is useful for treating airway  
 CC hyperresponsiveness or pulmonary inflammation, which is associated with  
 CC asthma, by inhibiting expression of human B7 protein. This sequence  
 CC represents an antisense oligonucleotide of the invention.  
 XX

SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGTGTCTCTGTCTAGTCCC 20

Db 1 CGTGTCTCTGTCTAGTCCC 20

#### RESULT 4

AA01045/c  
ID AA01045 standard; cDNA; 124 BP.

XX AC AA01045;

XX DT 07-MAY-1996 (first entry)

XX DE Human B7-2 exon h1A (signal).

XX KW T-cell costimulatory molecule; B7-2; T-lymphocyte; CD28; CTLA4; receptor;  
immunoglobulin; signal peptide; ds.  
XX OS Homo sapiens.

FH Key Location/Qualifiers  
FT CDS 107..124  
FT /\*tag= a

XX PN W09523859-A2.

XX PD 08-SEP-1995.

XX PF 02-MAR-1995; 95WO-US002576.

XX PR 02-MAR-1994; 94US-00205697.

XX PA (BGM ) BRIGHAM & WOMENS HOSPITAL.

XX PA (DAND ) DANA FARBER CANCER INST.

XX PI Sharpe AH, Borriello F, Freeman GJ, Nadler LM;

XX DR WPI: 1995-320574/41.

XX DR P-PSDB; AAR82898.

XX PT Novel T cell co-stimulatory molecules - corresponding to naturally  
occurring alternatively spliced forms of T cells co-stimulatory molecules  
or variants.  
XX PS Claim 40; Page 85; 111pp; English.

XX CC Human T-cell costimulatory molecule B7-2 gene exon h1A (AA01045) encodes  
the B7-2 signal peptide (AAR82898). Exon h1A can be utilised in the  
construction of nucleic acids used in the prodn. of soluble forms of T-  
cell costimulatory molecules that bind to CD28 or CTLA4 and trigger a  
costimulatory signal in T-cells

SQ Sequence 124 BP; 35 A; 30 C; 33 G; 26 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 20; DB 2; Length 124;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGTGTCTCTGTCTAGTCCC 20

Db 57 CGTGTCTCTGTCTAGTCCC 38

#### RESULT 5

AA00427/c  
ID AAC00427 standard; cDNA; 430 BP.

XX

AC AAC00427;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 425.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-00200610.

XX PR 26-FEB-1999; 99US-0122487P.

XX PA (GEST ) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX DR P-PSDB; AAG00421.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX PS Claim 1; SEQ ID NO 425; 71pp + Sequence Listing; English.

XX CC The present sequence is one of a large number of 5' ESTs derived from  
mRNAs encoding secreted proteins. An ORF has been identified within the  
sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
derived from 30 different tissues. EST sequences usually correspond  
mainly to the 3' untranslated region (UTR) of the mRNA because they are  
often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
well suited for isolating cDNA sequences derived from the 5' ends of  
mRNAs and even in those cases where longer cDNA sequences have been  
obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
mRNAs with intact 5' ends and can therefore be used to obtain full length  
cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
gene therapy and chromosome mapping procedures. They are used to obtain  
upstream regulatory sequences and to design expression and secretion  
vectors

SQ Sequence 430 BP; 117 A; 94 C; 104 G; 113 T; 0 U; 2 Other;

Query Match

Best Local Similarity 100.0%; Score 20; DB 3; Length 430;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGTGTGTCTGTGCTAGTCCC 20

Db 75 CGTGTGTCTGTGCTAGTCCC 56

#### RESULT 6

AA081351/c

ID AA081351 standard; cDNA; 1120 BP.

XX AC AA081351;

XX DT 25-MAR-2003 (revised)

XX DT 20-AUG-1995 (first entry)

XX DE Human B lymphocyte antigen B7-2 (hB7-2-clone 29).

XX KW CTLA4/CD28; counter receptor; B lymphocyte antigen; B7-2; ss.

XX OS Homo sapiens.

```

FH Key      Location/Qualifiers
FT CDS      107..1093
FT          /*tag= a
EN
PN
XX WO9503408-A1.
XX
XX 02-FEB-1995.
XX
XX 26-JUL-1994; 94WO-US008423.
XX
XX 26-JUL-1993; 93US-00101624.
XX
XX 19-AUG-1993; 93US-00109393.
XX
XX 03-NOV-1993; 93US-00147773.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX (REPK ) REPLIGEN CORP.
XX
XX Freeman GJ, Nadler LM, Gray GS, Greenfield E;
XX
XX WPI; 1995-075236/10.
XX P-PSDB; AAR67984.
XX
XX Nucleic acids encoding CTLA4/CD28 counter receptor, B7-2 - useful for
XX enhancing or suppressing T-cell mediated immune responses.
XX
XX Claim 4; Fig 8; 175pp; English.
XX
XX A cDNA library was constructed in the pCDM8 vector using poly A+ RNA from
XX the human anti-IgM activated B cells. Four clones were strongly positive
XX for B7-2 expression by indirect immunofluorescence using CTLA4lg and flow
XX cytometric analysis. The B7-2 cDNA insert in clone 29 was sequenced in
XX the pCDM8 expression vector employing the following strategy. Initial
XX sequencing was performed using sequencing primers T7 (AAQ81352), CDM8R
XX (AAQ81353) (Invitrogen) homologous to pCDM8 vector sequences adjacent to
XX the clone B7-2 cDNA. Sequencing was performed using dye terminator
XX chemistry and an ABI automated DNA sequencer. DNA sequence obtd. using
XX these primers was used to design additional sequencing primers (see
XX AAQ81354-Q81363). This cycle of sequencing and selection of additional
XX primers was continued until the B7-2 cDNA was completely sequenced on
XX both strands. The human B7-2 clone 29 cDNA sequence is given in AAQ81351.
XX The predicted protein sequence (AAR67984) exhibits many features common
XX to other type 1 Ig superfamily membrane proteins. Following cleavage of
XX the signal peptide the resulting membrane-bound protein would have an
XX unmodified mol. wt. of approx. 34 kDa. The extracellular domain contains
XX eight potential N-linked glycosylation sites. E. coli transfected with a
XX vector contg. the cDNA insert of clone 29 was deposited under ATCC 69357
XX on July 26 1993. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 20; DB 2; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTCGTGCTAGTCCC 20
Db |||||
57 CGTGTGTCGTGCTAGTCCC 38

RESULT 7
AAT49181/c
ID AAT49181 standard; cDNA; 1120 BP.
XX
XX AAT49181;
AC
XX
XX 08-APR-1997 (first entry)
DT
XX
XX Human B lymphocyte antigen B7-2 cDNA.
DE
XX
XX CTLA4; CD28; ligand; B7-2; B lymphocyte antigen; B-cell; costimulation;
KW immunoglobulin; antibody; autoimmune disease; allergy; tumour; vaccine;
KW graft versus host disease; T-cell; T lymphocyte; TH2 response;
KW immunosuppressive; immunostimulant; therapy; ss.

FH Key      Location/Qualifiers
FT CDS      107..1093
FT          /*tag= a
EN
PN
XX WO9640915-A2.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-US009052.
XX
XX 07-JUN-1995; 95US-00479744.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX (REPK ) REPLIGEN CORP.
XX
XX Freeman GJ, Nadler LM, Gray GS;
XX
XX WPI; 1997-077269/07.
XX P-PSDB; AAW08467.
XX
XX DNA encoding a B7-2 fusion protein - used to enhance or down regulate B
XX lymphocyte antigens.
XX
XX Example 4; Page 93-94; 171pp; English.
XX
XX A cDNA clone (AAT49181), designated clone 29, codes for the B-lymphocyte
XX antigen B7-2 (AAW08467), a CTLA4/CD28 ligand which costimulates T cell
XX activation. It was obtd. by transfecting COS cells with a human anti-IgM
XX activated B cell cDNA library, reacting transfectants with CTLA4lg and
XX CD28lg, and panning with anti-human IgG antibody. E. coli transfected
XX with a vector contg. the cDNA insert of clone 29 was deposited as ATCC
XX 69357. Nucleic acids encoding the extracellular domain, variable region-
XX like domain or constant region-like domain of B7-2 (see also AAT49197-98)
XX are used to construct novel fusion proteins with e.g. an immunoglobulin
XX constant region. These can be expressed in host cells and used to enhance
XX or suppress T cell-mediated immune responses
XX
XX Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 20; DB 2; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTCGTGCTAGTCCC 20
Db |||||
57 CGTGTGTCGTGCTAGTCCC 38

RESULT 8
AAV55784/c
ID AAV55784 standard; cDNA; 1120 BP.
XX
XX AAV55784;
AC
XX
XX 23-MAR-1999 (first entry)
DT
XX
XX Human B7-2 antigen coding sequence.
DE
XX
XX B7-2 antigen; mammalian tumour cell; T cell costimulation; CD28 ligand;
KW CTLA4 ligand; therapy; T-cell response; human; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX Key      Location/Qualifiers
FT CDS      107..1096
FT          /*tag= a

```

XX US5861310-A.  
 PN  
 XX  
 PD 19-JAN-1999.  
 XX  
 PF 30-MAY-1995; 95US-00456104.  
 XX  
 PR 03-NOV-1993; 93US-00147773.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Gray GS, Freeman GJ, Nadler LM;  
 XX  
 DR WPI; 1999-130394/11.  
 XX  
 PR P-PSDB; AAW73638.  
 XX  
 PT Tumour cell transfected to express B7-2 molecule - useful for tumour  
 PT therapy by stimulating T-cell response.  
 XX  
 PS Claim 9; Col 27-30; 27pp; English.  
 XX  
 CC This sequence encodes the human B7-2 antigen, which can be used in the  
 CC method of the invention. The method is for transfecting an isolated  
 CC mammalian tumour cell with an exogenous nucleic acid molecule encoding a  
 CC mammalian B7-2 molecule, where the B7-2 molecule is expressed in the  
 CC tumour cell is capable of costimulating a T cell and is capable of  
 CC binding a CD28 or CTLA4 ligand. The method is useful for treating tumours  
 CC by stimulating a T-cell response against tumour cells in vivo  
 XX  
 SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 2; Length 1120;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGTGTCTCTGTGCTAGTCCC 20  
 |||||  
 Db 57 CGTGTCTCTGTGCTAGTCCC 38  
 RESULT 9  
 AAC84049/c  
 ID AAC84049 standard; cDNA; 1120 BP.  
 XX  
 AC AAC84049;  
 XX  
 DT 28-MAR-2001 (first entry)  
 XX  
 DE Human B lymphocyte antigen B7-2 cDNA clone 29.  
 XX  
 KW Immunomodulator; fusion protein; human; murine; mouse; lymphocyte; CD28;  
 KW antigen; extracellular domain; CTLA4; immunoglobulin constant region;  
 KW immunogenicity; tumour; sarcoma; antigen presenting cell; macrophage;  
 KW T cell-mediated immune response; transplantation; vaccination; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6130316-A.  
 XX  
 PD 10-OCT-2000.  
 XX  
 PP 26-JUL-1994; 94US-00280757.  
 XX  
 PR 26-JUL-1993; 93US-00101624.  
 PR 19-AUG-1993; 93US-00109393.  
 PR 03-NOV-1993; 93US-00147773.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 PA (REPK ) REPLIGEN CORP.  
 XX  
 PI Freeman GJ, Nadler LM, Gray GS, Greenfield B;  
 XX  
 DR WPI; 2000-655681/63.

DR P-PSDB; AAB37085.  
 XX  
 PT Nucleic acids and fusion proteins of CTLA4/CD28 ligands, useful for  
 PT enhancing or suppressing T cell-mediated immune responses, especially  
 PT during tissue, skin or organ transplantation, or in graft-versus-host  
 PT disease.  
 XX  
 PS Claim 43; Fig 8; 83pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC fusion protein comprising a first nucleotide sequence encoding a first  
 CC peptide, and a second nucleotide sequence encoding a second peptide. The  
 CC first nucleotide sequence hybridizes in 6 x sodium chloride/sodium  
 CC citrate (SSC) at 45 deg. C, followed by a wash in 0.2 x SSC at 50 deg. C  
 CC to a portion of a nucleotide sequence which encodes a human or murine B  
 CC lymphocyte antigen (B7-2) extracellular domain. The first peptide has the  
 CC ability to bind CD28 or CTLA4. The first peptide has an amino acid  
 CC sequence that is identical or at least 50% identical with the  
 CC extracellular domain of a human B7-2 peptide (AAB37085). The second  
 CC peptide is especially an immunoglobulin constant region. This sequence  
 CC represents the cDNA clone 29 encoding the human B lymphocyte antigen B7-2  
 CC (hB7-2 clone 29) and is used as a first sequence in the construct of the  
 CC invention. The nucleic acid molecules are useful in various expression  
 CC vectors to direct synthesis of the corresponding proteins or peptides in  
 CC a variety of hosts, particularly eukaryotic cells, e.g. mammalian or  
 CC insect cell culture. The nucleic acids are also useful for enhancing the  
 CC immunogenicity of a mammalian cell, e.g. tumour cell (sarcoma) or an  
 CC antigen presenting cell (macrophage). The fusion proteins or peptides are  
 CC useful for enhancing or suppressing T cell-mediated immune responses,  
 CC e.g. in situations of tissue, skin or organ transplantation, or in graft-  
 CC versus-host disease. The proteins are also useful for enhancing the  
 CC efficacy of vaccination against a variety of pathogens, and may also be  
 CC used to upregulate an immune response against a particular pathogen  
 CC during an infection or against a tumour in a tumour-bearing host  
 XX  
 SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 3; Length 1120;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGTGTCTCTGTGCTAGTCCC 20  
 |||||  
 Db 57 CGTGTCTCTGTGCTAGTCCC 38  
 RESULT 10  
 ABV72340/c  
 ID ABV72340 standard; cDNA; 1120 BP.  
 XX  
 AC ABV72340;  
 XX  
 DT 16-DEC-2002 (first entry)  
 XX  
 DE Nucleotide sequence of human B7-2 (CD86).  
 XX  
 KW B7 protein; B7-2; CD86; CD28 ligand; T cell; T cell proliferation;  
 KW infectious disease; cancer; immunotherapy; immunotherapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 107..1096  
 FT /\*tag= a  
 FT /product= "B7-2"  
 XX  
 PN US2002115214-A1.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 26-JAN-1996; 96US-00592711.  
 XX  
 PR 23-NOV-1988; 88US-00275433.

PR 07-APR-1992; 92US-00864805.  
 PR 07-APR-1992; 92US-00864807.  
 PR 07-APR-1992; 92US-00864866.  
 PR 04-JUN-1993; 93US-00073223.  
 PR 03-JUN-1994; 94US-00253964.  
 PR 10-MAR-1995; 95US-00403253.  
 PR 04-MAY-1995; 95US-00435816.  
 XX (JUNE/) JUNE C H.  
 PA (THOM/) THOMPSON C B.  
 PA (NABEL/) NABEL G J.  
 PA (GRAY/) GRAY G S.  
 PA (RENN/) RENNETT P D.  
 XX  
 PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;  
 XX  
 DR WPI: 2002-712476/77.  
 DR P-PSDB; ABB78364.  
 XX  
 PT Inducing a population of T cells to proliferate, by activating population  
 PT of T cells and stimulating an accessory molecule on the surface of the T  
 PT cells with a ligand which binds the accessory molecule.  
 XX  
 PS Disclosure; Page 41-43; 88pp; English.  
 XX  
 CC The present sequence encodes a member of the B7 family of protein, B7-2  
 CC (CD86). B7 proteins are ligands for CD28. Activated T cells are contacted  
 CC with a stimulatory form of a natural ligand for CD28, such as a B7  
 CC protein, for costimulation. B7 molecules are used in the method of the  
 CC invention. The specification describes method for inducing a population  
 CC of T cells to proliferate. The method involves activating population of T  
 CC cells, stimulating an accessory molecule (e.g. CD28) on T cell surface  
 CC with a ligand (e.g. B7 protein) which binds the molecule, to induce  
 CC proliferation of T cells, monitoring proliferation of T cells in response  
 CC to continuing exposure to the ligand, and reactivating and restimulating  
 CC T cells when rate of proliferation has decreased to induce further  
 CC proliferation of the cells. The method is useful for inducing  
 CC cancer and immunotherapy. The method allows for the expansion of a  
 CC population of T cells in numbers sufficient to reconstitute an  
 CC individual's total CD4+ or CD8+ T cell population. The resulting T cell  
 CC population can be genetically transduced and used for immunotherapy or  
 CC can be used in methods of in vitro analyses of infectious agents. A  
 CC population of tumour-infiltrating lymphocytes can be obtained from an  
 CC individual afflicted with cancer and the T cells stimulated to  
 CC proliferate to sufficient numbers. The resulting T cell population can be  
 CC genetically transduced to express tumour necrosis factor (TNF) or other  
 CC factor and restored to the individual. CD4+ T cells expanded by this  
 CC method are useful in the treatment of HIV infection in an individual  
 XX  
 SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 6; Length 1120;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGTGTCTGTCTAGTCC 20  
 DB 57 CGTGTCTGTCTAGTCC 38  
 RESULT 11  
 AAD27968/c  
 ID AAD27968 standard; cDNA; 1120 BP.  
 XX  
 AC AAD27968;  
 XX  
 DT 16-JUL-2002 (first entry)  
 XX  
 DE Human B7-2 cDNA.  
 XX  
 KW T cell; CD3; accessory molecule; CD28; cancer; infectious disease;  
 KW immunotherapy; human immunodeficiency virus; HIV infection; cytokine;

KW human; B7-2; CD86; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 107..1096  
 FT /\*tag= a  
 FT /product= "B7-2 protein"  
 XX  
 PN US6352694-B1.  
 XX  
 PD 05-MAR-2002.  
 XX  
 PF 10-MAR-1995; 95US-00403253.  
 XX  
 PR 03-JUN-1994; 94US-00253964.  
 XX  
 PA (GENY ) GENETICS INST INC.  
 PA (UNMI ) UNIV MICHIGAN.  
 XX  
 PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;  
 XX  
 DR WPI: 2002-314696/35.  
 DR P-PSDB; AAE14634.  
 XX  
 PT Inducing T cell population to proliferate, useful in cancer therapy,  
 PT comprises activating T cells by contacting T cells in vitro with  
 PT immobilized anti-CD3 antibody and stimulating accessory molecule on T  
 PT cell surface.  
 XX  
 PS Example 11; Col 61-64; 71pp; English.  
 XX  
 CC The invention relates to a method of inducing T cell population to  
 CC proliferate for use in therapy comprising activating T cells by  
 CC contacting T cells in vitro with anti-CD3 antibody which is immobilised  
 CC on solid phase surface, and stimulating accessory molecule on T cell  
 CC surface in vitro with anti-CD28 antibody, or stimulatory form of natural  
 CC ligand for CD28 such as B7-1 or B7-2. The method is useful for inducing a  
 CC population of T cells to proliferate in sufficient numbers for use in  
 CC therapy e.g., for treating cancer or an infectious disease. The method  
 CC can be used to selectively expand the population of CD28+, CD4+, CD8+,  
 CC CD28RA+ or CD28RO+ T cells for immunotherapy. The T cell population  
 CC resulting by the method can be genetically transduced and used for  
 CC immunotherapy or can be used for in vitro analysis of infectious agents  
 CC such as human immunodeficiency virus (HIV). Proliferation of a population  
 CC of CD4+ T cells obtained from an individual infected with HIV can be  
 CC achieved and the cells rendered resistant to HIV infection. Following the  
 CC expansion of the T cells to sufficient numbers, the expanded T cells are  
 CC restored to the individual. Also CD4+ T cells expanded by the above  
 CC mentioned is useful for treating HIV infection in an individual. A  
 CC population of tumour-infiltrating lymphocytes can be obtained from an  
 CC individual afflicted with cancer and the T cells stimulated to  
 CC proliferate to sufficient numbers and restored to the individual. The  
 CC supernatants from cultures of T cells expanded from above mentioned  
 CC method are useful as a rich source of cytokines and can be used to  
 CC sustain T cells in vivo or ex vivo. Stimulating and expanding a  
 CC population of antigen specific T cells are useful in therapeutic  
 CC conditions where it is desirable to upregulate an immune response. The T  
 CC cell proliferation occurs in the absence of exogenous growth factors or  
 CC accessory cells. The present sequence is human B7-2 (CD86) cDNA used in  
 CC the invention  
 XX  
 SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 6; Length 1120;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGTGTCTGTCTAGTCC 20  
 DB 57 CGTGTCTGTCTAGTCC 38

RESULT 12  
AAD60974/c  
ID AAD60974 standard; cDNA; 1120 BP.  
XX  
AC AAD60974;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human CD28 ligand, B7-2 (CD86) cDNA.  
XX  
KW Human; HIV infection; human immunodeficiency virus; CD28 ligand; B7;  
XX therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 107..1096  
FT /\*tag= a  
FT /product= "CD28 ligand, B7-2 protein"  
XX  
PN US2003099643-A1.  
XX  
PD 29-MAY-2003.  
XX  
PF 08-JUL-1999; 99US-00350202.  
XX  
PR 23-NOV-1988; 88US-00275433.  
PR 22-NOV-1989; 89WO-US005304.  
PR 07-APR-1992; 92US-00864805.  
PR 07-APR-1992; 92US-00864807.  
PR 07-APR-1992; 92US-00864866.  
PR 04-JUN-1993; 93US-00073223.  
PR 03-JUN-1994; 94US-00253694.  
PR 10-MAR-1995; 95US-00403253.  
XX  
PA (JUNE/) JUNE C H.  
PA (THOM/) THOMPSON C B.  
PA (NABEL/) NABEL G J.  
PA (GRAY/) GRAY G S.  
PA (RENN/) RENNERT P D.  
XX  
PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;  
XX  
DR WPI; 2003-801206/75.  
DR P-PSDB; ABW00432.  
XX  
PT Treating HIV infection in individual by isolating T cells from  
PT leukocytes, contacting T cells with anti-CD3 antibody for T cell  
PT proliferation, separating antibody from T cells, monitoring proliferation  
PT of T cells.  
XX  
PS Disclosure; Page 34-35; Opp; English.  
XX  
CC The present invention relates to a novel method of treating human  
CC immunodeficiency virus (HIV) infection in an individual. The method  
CC involves isolating population of CD4 T cells from leukocytes, contacting  
CC population of CD4+ T cells with an anti-CD3 antibody for stimulating T  
CC cell proliferation, separating antibody from T cells, monitoring  
CC proliferation of T cells, restimulating T cells with antibody and  
CC restoring T cells to individual. The present sequence is human CD28  
CC ligand, B7 cDNA. This sequence is used to illustrate the method of the  
CC invention  
XX  
SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20  
|||||  
DB 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 13  
ADK69859/c  
ID ADK69859 standard; cDNA; 1120 BP.  
XX  
AC ADK69859;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human B7-2 cDNA.  
XX  
KW CD28-associated signal; immunotherapy; infectious disease; cancer;  
XX leukopheresis; human; B7-2; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 107..1093  
FT /\*tag= a  
FT /product= "Human B7-2 protein"  
XX  
PN US6534055-B1.  
XX  
PD 18-MAR-2003.  
XX  
PF 04-MAY-1995; 95US-00435816.  
XX  
PR 07-APR-1992; 92US-00864805.  
PR 07-APR-1992; 92US-00864807.  
PR 07-APR-1992; 92US-00864866.  
PR 04-JUN-1993; 93US-00073223.  
PR 03-JUN-1994; 94US-00253964.  
PR 10-MAR-1995; 95US-00403253.  
XX  
PA (GEMY ) GENETICS INST INC.

XX  
PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;  
XX  
DR WPI; 2003-531074/50.  
DR P-PSDB; ADK69860.  
XX  
PT Expanding T cell populations, useful for preparing renewable sources of T  
PT cells (e.g. CD4+) for immunotherapy or diagnostics, by stimulating a CD28  
PT -associated signal on the surface of the cells with an anti-CD28  
PT antibody, B7-1 or B7-2.  
XX  
PS Disclosure; SEQ ID NO 3; 82pp; English.  
XX  
CC The invention relates to a method for expanding a population of T cells  
CC to about 100-100000-fold over the original T cell population, or to about  
CC 10 log 1.0 to 12 log 1.2. The method comprises stimulating a CD28-  
CC associated signal on the surface of the T cells with agent comprising an  
CC anti-CD28 antibody, B7-1 or B7-2. The method is useful for inducing a  
CC population of T cells (e.g. CD4+, CD28+, CD8+, CD28RA+ or CD28RO+ T  
CC cells) to proliferate. The method is particularly useful for preparing a  
CC renewable source of CD4+ T cells. The expanded T cell population can be  
CC genetically transduced, and used for immunotherapy to treat a variety of  
CC human diseases (e.g. infectious diseases or cancer), or used in  
CC diagnostic protocols. T cells were obtained from leukopheresis of a  
CC normal donor, and purified with FICOLL density gradient centrifugation,  
CC followed by magnetic immunobead sorting. The present sequence is human B7  
CC -2 cDNA.  
XX  
SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20  
|||||  
DB 57 CGTGTGCTGTGCTAGTCCC 38



RESULT 14  
ADI19363/c  
ID ADI19363 standard; cDNA; 1120 BP.

XX AC ADI19363;  
XX DT 15-APR-2004 (first entry)  
XX DE Human B7-2 cDNA, CD86.  
XX KW T cell; ss; immunotherapy; therapy; HIV infection; cancer;  
XX KW infectious disease; cytostatic; antimicrobial; gene; human; chromosome 3.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX CDS 107..1096  
XX FT /\*tag= a  
XX FT /product= "Human B7-2 protein"  
XX PN US2004001829-A1.  
XX PD 01-JAN-2004.

XX PF 17-MAR-2003; 2003US-00390330.  
XX PR 22-NOV-1988; 88US-00275433.  
XX PR 23-NOV-1989; 89WO-US005304.  
XX PR 07-APR-1992; 92US-00864805.  
XX PR 07-APR-1992; 92US-00864807.  
XX PR 07-APR-1992; 92US-00864866.  
XX PR 04-JUN-1993; 93US-00073223.  
XX PR 03-JUN-1994; 94US-00253164.  
XX PR 10-MAR-1995; 95US-00403253.  
XX PR 04-MAY-1995; 95US-00435816.

XX PA (JUNE/) JUNE C H.  
XX PA (THOM/) THOMPSON C B.  
XX PA (NABEL/) NABEL G J.  
XX PA (GRAY/) GRAY G S.  
XX PA (RENN/) RENNETT P D.

XX PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennett PD;  
XX DR WPI; 2004-061648/06.  
XX DR P-PSDB; ADI19364.

XX PT Inducing a population of T cells to proliferate, for immunotherapy or  
XX PT treating HIV infection, cancer or infectious disease, comprises  
XX PT activating a population of T cells and stimulating an accessory molecule  
XX PT on the surface of the T cells.

XX PS Disclosure; SEQ ID NO 3; 80pp; English.

XX CC The invention relates to a method for inducing a population of T cells to  
XX CC proliferate. The method comprising activating a population of T cells,  
XX CC and stimulating an accessory molecule on the surface of the T cells with  
XX CC a ligand that binds the accessory molecule. The invention is useful for  
XX CC immunotherapy, for treating HIV infection, cancer or infectious disease,  
XX CC or in diagnostic applications. The present sequence is human B7-2 cDNA.  
XX CC B7 gene is located on chromosome 3.

XX SQ Sequence 1120 BP; 354 A; 238 C; 229 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20.  
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 15  
ADJ54476/c  
ID ADJ54476 standard; DNA; 2781 BP.  
XX AC ADJ54476;  
XX DT 06-MAY-2004 (first entry)  
XX DE Human B7-2 DNA #2.  
XX KW Airway hyperresponsiveness; pulmonary inflammation;  
XX KW antisense oligonucleotide; human; B7 protein; B7-2; asthma;  
XX KW antiasthmatic; antiinflammatory; gene; ds.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX CDS 117..1106  
XX FT /\*tag= a  
XX FT /product= "Human B7-2 #2"

XX PN US2004023917-A1.

XX PD 05-FEB-2004.

XX PF 23-MAY-2003; 2003US-00444206.

XX PR 31-DEC-1996; 96US-00777266.

XX PR 04-JUN-1999; 99US-00326186.

XX PR 25-MAY-2000; 2000WO-US014471.

XX PR 09-MAY-2001; 2001US-00851871.

XX PA (BENN/) BENNETT C F.

XX PA (VICK/) VICKERS T A.

XX PA (KARR/) KARRAS J G.

XX PI Bennett CF, Vickers TA, Karras JG;

XX DR WPI; 2004-132608/13.  
XX DR P-PSDB; ADJ54626.  
XX DR GENBANK; BC040261.

XX PT Treating airway hyperresponsiveness or pulmonary inflammation comprises  
XX PT administering an antisense compound targeted to a nucleic acid molecule  
XX PT encoding a human B7 protein to the individual.

XX PS Example 27; SEQ ID NO 296; 182pp; English.

XX CC The invention relates to a method for treating airway hyperresponsiveness  
XX CC or pulmonary inflammation in an individual comprising administering an  
XX CC antisense compound targeted to a nucleic acid molecule encoding a human  
XX CC B7 protein. The invention also relates to a method of inhibiting  
XX CC expression of a nucleic acid molecule encoding B7-1 or B7-2. The  
XX CC antisense compound is an antisense oligonucleotide which has a modified  
XX CC sugar moiety and nucleobase. The human B7 protein is human B7-1 or B7-2  
XX CC protein or both. The compound is useful for treating airway  
XX CC hyperresponsiveness or pulmonary inflammation, which is associated with  
XX CC asthma, by inhibiting expression of human B7 protein. This sequence  
XX CC represents DNA encoding a human B7-2 polypeptide of the invention.

XX SQ Sequence 2781 BP; 887 A; 575 C; 597 G; 722 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 2781;

Best Local Similarity 100.0%; Pred. No. 4.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20  
Db 67 CGTGTGCTGTGCTAGTCCC 48

RESULT 16

```

ADP10461/c
ID ADP10461 standard; DNA; 2781 BP.
XX
AC ADP10461;
XX
DT 12-AUG-2004 (first entry)
XX
DE Reference mRNA sequences for marker probe #138.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
XX
OS Homo sapiens.
XX
FN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 80; SEQ ID NO 470; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC and monitoring of allograft rejection and other disorders.
XX
SQ Sequence 2781 BP; 886 A; 575 C; 598 G; 722 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 2781;
Best Local Similarity 100.0%; Pred. No. 4.6; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

Qy 1 CGTGTCTCTGCTAGTCC 20
Db 67 CGTGTCTCTGCTAGTCC 48

RESULT 17
ADJ54477/c
ID ADJ54477 standard; DNA; 68001 BP.
XX
AC ADJ54477;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human B7-2 genomic DNA.
XX

Airway hyperresponsiveness; pulmonary inflammation;
antisense oligonucleotide; human; B7 protein; B7-2; asthma;
antiasthmatic; antiinflammatory; gene; ds.
Homo sapiens.
US2004023917-A1.
05-FEB-2004.
23-MAY-2003; 2003US-0044206.
31-DEC-1996; 96US-00777266.
04-JUN-1999; 99US-00326186.
25-MAY-2000; 2000WO-US014471.
09-MAY-2001; 2001US-00851871.
(BENN/) BENNETT C F.
(VICK/) VICKERS T A.
(KARR/) KARRAS J G.
Bennett CF, Vickers TA, Karras JG;
WPI; 2004-132608/13.
GENBANK; NT_005543.
Treating airway hyperresponsiveness or pulmonary inflammation comprises
administering an antisense compound targeted to a nucleic acid molecule
encoding a human B7 protein to the individual.
Example 27; SEQ ID NO 297; 182pp; English.
The invention relates to a method for treating airway hyperresponsiveness
or pulmonary inflammation in an individual comprising administering an
antisense compound targeted to a nucleic acid molecule encoding a human
B7 protein. The invention also relates to a method of inhibiting
expression of a nucleic acid molecule encoding B7-1 or B7-2. The
antisense compound is an antisense oligonucleotide which has a modified
sugar moiety and nucleobase. The human B7 protein is human B7-1 or B7-2
protein or both. The compound is useful for treating airway
hyperresponsiveness or pulmonary inflammation, which is associated with
asthma, by inhibiting expression of human B7 protein. This sequence
represents human B7-2 genomic DNA of the invention.
Sequence 68001 BP; 20352 A; 13494 C; 14117 G; 20038 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 68001;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTCTGCTAGTCC 20
Db 1067 CGTGTCTCTGCTAGTCC 1048

RESULT 18
ADN72456
ID ADN72456 standard; cDNA; 3168 BP.
XX
AC ADN72456;
XX
DT 15-JUL-2004 (first entry)
XX
DE Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 351.
KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX
FN WO2004035798-A2.
XX

```

PD 29-APR-2004.  
XX  
XX  
PF 20-OCT-2003; 2003WO-EP011658.  
XX  
XX  
PR 18-OCT-2002; 2002EP-00079408.  
XX  
XX  
PA (CROP-) CROPDESIGN NV.  
XX  
XX  
PI Inze D, De Veylder L, Vlieghe K;  
XX  
XX  
DR WPI: 2004-348466/32.  
DR P-PSDB; ADN72457.  
XX  
XX  
PT Altering plant characteristics, useful for producing plants for enzyme or  
PT pharmaceutical production comprising modifying in a plant, expression of  
PT one or more nucleic acids and/or modifying level or activity of one or  
PT more proteins.  
XX  
XX  
PS Claim 1; SEQ ID NO 351; 134pp; English.  
XX  
XX  
CC This invention relates to a novel method for altering one or more plant  
CC characteristics. Specifically, it refers to identifying genes that are up  
CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
CC alter plant characteristics accordingly. The present invention describes  
CC generating transgenic plants for the production of growth regulators,  
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
CC the altered plant characteristics are selected from increased yield or  
CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
CC or physiology, altered endoreplication, biochemistry, signal  
CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
CC each relative to the corresponding wild type plants. Accordingly, these  
CC sequences can also be useful as positive or negative selectable markers  
CC during transformation of cells or tissues. The identified genes play a  
CC role in a variety of biological processes such as DNA replication, cell  
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
CC transcription factors. This polynucleotide sequence is the cDNA  
CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa  
CC transcription factor, given in an exemplification of the invention.  
XX  
XX  
SQ Sequence 3168 BP; 1042 A; 542 C; 760 G; 724 T; 0 U; 0 Other;  
Query Match 87.0%; Score 17.4; DB 12; Length 3168;  
Best Local Similarity 94.7%; Pred. No. 95;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GTGTCTCTGCTAGTCC 20  
Db 2429 GTGTCTCTGCTAGTCC 2447  
RESULT 19  
AAC05026  
ID AAC05026 standard; cDNA; 368 BP.  
XX  
XX  
AC AAC05026;  
XX  
XX  
DT 06-OCT-2000 (first entry)  
XX  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 9101.  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
FN EP1033401-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
XX  
PR 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.  
XX  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX  
XX  
DR WPI: 2000-500381/45.  
XX  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
XX  
PS Claim 1; SEQ ID NO 9101; 71pp + Sequence Listing; English.  
XX  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dr primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors  
XX  
XX  
SQ Sequence 368 BP; 79 A; 122 C; 59 G; 105 T; 0 U; 3 Other;  
Query Match 80.0%; Score 16; DB 3; Length 368;  
Best Local Similarity 88.9%; Pred. No. 3.6e+02;  
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GTGTCTCTGCTAGTCC 19  
Db 192 GTGTCTCTGCTAGTCC 209  
RESULT 20  
AAL01062  
ID AAL01062 standard; cDNA; 372 BP.  
XX  
XX  
AC AAL01062;  
XX  
XX  
DT 21-NOV-2001 (first entry)  
XX  
XX  
DE Human reproductive system related antigen cDNA SEQ ID NO: 1063.  
KW Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ss.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
FN WO200155320-A2.  
XX  
XX  
PD 02-AUG-2001.  
XX  
XX  
PF 17-JAN-2001; 2001WO-US001339.  
XX  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225211P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-02331968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465570/50.

XX P-PSDB; AAM95092.

PT Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.

PS Claim 1; SEQ ID NO 1063; 1297pp + Sequence Listing; English.

XX The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a coding sequence of the CC invention

XX Sequence 372 BP; 104 A; 68 C; 96 G; 101 T; 0 U; 3 Other;

Query Match 80.0%; Score 16; DB 4; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;

Matches	16;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	2	GTGTGCTGTGCTAGT	17						
Db	70	GTGTGCTGTGCTAGT	85						
RESULT 21									
ABL96529									
ID	ABL96529 standard; cDNA; 372 BP.								
XX	AC	ABL96529;							
DT	21-JUN-2002 (first entry)								
XX	Human testicular antigen encoding cDNA SEQ ID NO: 197.								
XX	Human; testicular antigen; testes; cancer; metastasis; immune disorder;								
KW	reproductive system disorder; urinary system disorder; gene therapy;								
KW	cardiovascular disorder; respiratory disorder; neurological disorder;								
KW	gastrointestinal disease; infection; cytostatic; gene; ss.								
XX									
OS	Homo sapiens.								
PN	WO200155317-A2.								
PD	02-AUG-2001.								
XX	17-JAN-2001; 2001WO-US001329.								
XX									
PR	31-JAN-2000;	2000US-0179065P.							
PR	04-FEB-2000;	2000US-0180628P.							
PR	24-FEB-2000;	2000US-0184664P.							
PR	02-MAR-2000;	2000US-0186350P.							
PR	16-MAR-2000;	2000US-0189874P.							
PR	17-MAR-2000;	2000US-0190076P.							
PR	18-APR-2000;	2000US-0198123P.							
PR	19-MAY-2000;	2000US-0205515P.							
PR	07-JUN-2000;	2000US-0209467P.							
PR	28-JUN-2000;	2000US-0214886P.							
PR	30-JUN-2000;	2000US-0215135P.							
PR	07-JUL-2000;	2000US-0216647P.							
PR	07-JUL-2000;	2000US-0216880P.							
PR	11-JUL-2000;	2000US-0217487P.							
PR	11-JUL-2000;	2000US-0217496P.							
PR	14-JUL-2000;	2000US-0218290P.							
PR	26-JUL-2000;	2000US-0220963P.							
PR	26-JUL-2000;	2000US-0220964P.							
PR	14-AUG-2000;	2000US-0224518P.							
PR	14-AUG-2000;	2000US-0224519P.							
PR	14-AUG-2000;	2000US-0225213P.							
PR	14-AUG-2000;	2000US-0225214P.							
PR	14-AUG-2000;	2000US-0225266P.							
PR	14-AUG-2000;	2000US-0225267P.							
PR	14-AUG-2000;	2000US-0225268P.							
PR	14-AUG-2000;	2000US-0225270P.							
PR	14-AUG-2000;	2000US-0225447P.							
PR	14-AUG-2000;	2000US-0225757P.							
PR	14-AUG-2000;	2000US-0225758P.							
PR	14-AUG-2000;	2000US-0225759P.							
PR	18-AUG-2000;	2000US-0226279P.							
PR	22-AUG-2000;	2000US-0226681P.							
PR	22-AUG-2000;	2000US-0226686P.							
PR	22-AUG-2000;	2000US-0227182P.							
PR	23-AUG-2000;	2000US-0227009P.							
PR	30-AUG-2000;	2000US-0228924P.							
PR	01-SEP-2000;	2000US-0229287P.							
PR	01-SEP-2000;	2000US-0229343P.							
PR	01-SEP-2000;	2000US-0229344P.							
PR	01-SEP-2000;	2000US-0229345P.							
PR	05-SEP-2000;	2000US-0229509P.							
PR	05-SEP-2000;	2000US-0229513P.							
PR	06-SEP-2000;	2000US-0230437P.							
PR	06-SEP-2000;	2000US-0230438P.							
PR	08-SEP-2000;	2000US-0231242P.							
PR	08-SEP-2000;	2000US-0231243P.							
PR	08-SEP-2000;	2000US-0231244P.							
PR	08-SEP-2000;	2000US-0231413P.							
PR	08-SEP-2000;	2000US-0231414P.							
PR	08-SEP-2000;	2000US-0232080P.							
PR	08-SEP-2000;	2000US-0232081P.							
PR	12-SEP-2000;	2000US-0231968P.							
PR	14-SEP-2000;	2000US-0232397P.							
PR	14-SEP-2000;	2000US-0232398P.							
PR	14-SEP-2000;	2000US-0232399P.							
PR	14-SEP-2000;	2000US-0232400P.							
PR	14-SEP-2000;	2000US-0232401P.							
PR	14-SEP-2000;	2000US-0233063P.							
PR	14-SEP-2000;	2000US-0233064P.							
PR	14-SEP-2000;	2000US-0233065P.							
PR	21-SEP-2000;	2000US-0234223P.							
PR	21-SEP-2000;	2000US-0234274P.							
PR	25-SEP-2000;	2000US-0234997P.							
PR	25-SEP-2000;	2000US-0234998P.							
PR	26-SEP-2000;	2000US-0235484P.							
PR	27-SEP-2000;	2000US-0235834P.							
PR	27-SEP-2000;	2000US-0235836P.							
PR	29-SEP-2000;	2000US-0236327P.							
PR	29-SEP-2000;	2000US-0236367P.							
PR	29-SEP-2000;	2000US-0236368P.							
PR	29-SEP-2000;	2000US-0236369P.							
PR	29-SEP-2000;	2000US-0236370P.							
PR	02-OCT-2000;	2000US-0236802P.							
PR	02-OCT-2000;	2000US-0237037P.							
PR	02-OCT-2000;	2000US-0237038P.							
PR	02-OCT-2000;	2000US-0237039P.							
PR	02-OCT-2000;	2000US-0237040P.							
PR	13-OCT-2000;	2000US-0239355P.							
PR	13-OCT-2000;	2000US-0239357P.							
PR	20-OCT-2000;	2000US-0240960P.							
PR	20-OCT-2000;	2000US-0241221P.							
PR	20-OCT-2000;	2000US-0241785P.							
PR	20-OCT-2000;	2000US-0241786P.							
PR	20-OCT-2000;	2000US-0241787P.							
PR	20-OCT-2000;	2000US-0241808P.							
PR	20-OCT-2000;	2000US-0241809P.							
PR	20-OCT-2000;	2000US-0241826P.							
PR	01-NOV-2000;	2000US-0244617P.							
PR	08-NOV-2000;	2000US-0246474P.							
PR	08-NOV-2000;	2000US-0246475P.							
PR	08-NOV-2000;	2000US-0246476P.							
PR	08-NOV-2000;	2000US-0246477P.							
PR	08-NOV-2000;	2000US-0246478P.							
PR	08-NOV-2000;	2000US-0246523P.							
PR	08-NOV-2000;	2000US-0246524P.							
PR	08-NOV-2000;	2000US-0246525P.							
PR	08-NOV-2000;	2000US-0246526P.							
PR	08-NOV-2000;	2000US-0246527P.							
PR	08-NOV-2000;	2000US-0246528P.							
PR	08-NOV-2000;	2000US-0246532P.							
PR	08-NOV-2000;	2000US-0246609P.							
PR	08-NOV-2000;	2000US-0246610P.							
PR	08-NOV-2000;	2000US-0246611P.							
PR	08-NOV-2000;	2000US-0246613P.							
PR	17-NOV-2000;	2000US-0249207P.							
PR	17-NOV-2000;	2000US-0249208P.							
PR	17-NOV-2000;	2000US-0249209P.							
PR	17-NOV-2000;	2000US-0249210P.							
PR	17-NOV-2000;	2000US-0249211P.							
PR	17-NOV-2000;	2000US-0249212P.							
PR	17-NOV-2000;	2000US-0249213P.							
PR	17-NOV-2000;	2000US-0249214P.							
PR	17-NOV-2000;	2000US-0249215P.							
PR	17-NOV-2000;	2000US-0249216P.							
PR	17-NOV-2000;	2000US-0249217P.							
PR	17-NOV-2000;	2000US-0249218P.							
PR	17-NOV-2000;	2000US-0249218P.							

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250191P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251473P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-02559678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-483232/52.  
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
PT for preventing, diagnosing and/or treating testicular cancer.  
XX  
XX Claim 1; SEQ ID NO 197; 766pp; English.  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a cDNA of the  
CC invention  
XX  
XX Sequence 372 BP; 104 A; 68 C; 96 G; 101 T; 0 U; 3 Other;  
SQ Query Match 80.0%; Score 16; DB 4; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 GTGTGCTGTGCTAGT 17  
DB 70 GTGTGCTGTGCTAGT 85  
  
RESULT 22  
ABX91230  
ID ABX91230 standard; cDNA; 515 BP.  
XX  
XX ABX91230;  
XX  
XX 07-MAY-2003 (first entry)  
XX Murine gene trapped sequence (GTS) SEQ ID No 574.  
XX Murine; mouse; gene trap technology; gene trapped sequence; GTS;  
KW gene identification; functional genomic analysis; gene discovery;  
KW gene expression analysis; cross species hybridisation analysis;  
KW antisense inhibition; gene targeting; gene; ss.  
XX  
OS Mus sp.  
XX  
XX US2002161207-A1.  
XX  
XX 31-OCT-2002.  
XX  
XX 30-NOV-2000; 2000US-00728444.  
XX

XX 01-DEC-1999; 99US-0168360P.  
XX (FRIE/) FRIEDRICH G.  
PA (ZAMB/) ZAMBROWICZ B.  
PA (SAND/) SANDS A T.  
XX Friedrich G, Zambrowicz B, Sands AT;  
PI WPI; 2003-288124/28.  
XX New murine polynucleotides comprising gene trapped sequences, useful in  
PT functional genomic analysis, in the development of new therapeutic or  
PT diagnostic agents, for diagnostic gene expression analysis or for genetic  
PT manipulations.  
XX  
XX Claim 2; SEQ ID NO 574; 29pp; English.  
XX The present invention relates to novel murine cDNAs produced using gene  
CC trap technology. The OMNIBANK gene trapped sequences (GTSS) are  
CC individually identified novel genes, and are useful in functional genomic  
CC analysis, in the discovery and development of new therapeutic and  
CC diagnostic agents, for gene discovery, for diagnostic gene expression  
CC analysis, for cross species hybridisation analysis, and for genetic  
CC manipulations such as antisense inhibition or gene targeting. The  
CC polynucleotides of the invention are also useful for isolating cDNAs,  
CC genomic clones or full-length genes/polynucleotides, or their homologues,  
CC heterologues, paralogues or orthologues, that are capable of hybridising  
CC to one or more of the new murine polynucleotide sequences. The  
CC polynucleotides are also useful for identifying the coding regions of the  
CC murine genome, and as hybridisation probes. ABX90657-ABX91862 represent  
CC the murine GTSS of the invention. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/peipsDIBentry.html  
XX  
XX Sequence 515 BP; 152 A; 101 C; 109 G; 151 T; 0 U; 2 Other;  
SQ Query Match 80.0%; Score 16; DB 8; Length 515;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CGTGTGCTGTGCTAG 16  
DB 384 CGTGTGCTGTGCTAG 399  
  
RESULT 23  
ABZ21938  
ID ABZ21938 standard; cDNA; 2175 BP.  
XX  
XX ABZ21938;  
XX  
XX 28-MAR-2003 (first entry)  
XX Human CAK1 antigen 9.46 encoding cDNA SEQ ID NO:1.  
XX Human; CAK1 antigen 9.46; antigen; malignant tumour; inflammation;  
KW immunological disease; haemopathy; HIV infection; gene; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
PH 245..505  
FT CDS /\*tag= a  
FT /product= "CAK1 antigen 9.46"  
XX  
XX CNI352019-A.  
XX  
XX 05-JUN-2002.  
XX  
XX 10-NOV-2000; 2000CN-00127321.  
XX



CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g. osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention  
XX

SQ Sequence 439 BP; 68 A; 116 C; 180 G; 74 T; 0 U; 1 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 439;  
Best Local Similarity 89.5%; Pred. No. 4.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTCTGCTAGTCC 19  
|||||  
Db 39 CGTGTCTCTGCGAGGCC 57  
|||||

RESULT 26

AAK82836/c  
ID AAK82836 standard; DNA; 468 BP.

XX AAK82836;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37648.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.



PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-483426/52.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
PS Disclosure; SEQ ID NO 37648; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the

CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 468 BP; 93 A; 136 C; 148 G; 91 T; 0 U; 0 Other;  
Query Match 79.0%; Score 15.8; DB 4; Length 468;  
Best Local Similarity 89.5%; Pred No. 4.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GTGTGTCGTGCTAGTCCC 20  
Db 53 GGGTGTCTGGGCTAGTCCC 35  
RESULT 27  
AAAS1678/c  
ID AAS1678 standard; cDNA; 1345 BP.  
XX  
AC AAS1678;  
XX  
DT 16-APR-2003 (first entry)  
DE Human nucleic acid associated protein (NAAP)-4 encoding cDNA.  
XX  
KW Human; nucleic acid associated protein; NAAP; cancer; atherosclerosis;  
KW cell proliferative disorder; neurological disorder; Huntington's disease;  
KW epilepsy; stroke; immune disorder; acquired immune deficiency syndrome;  
KW AIDS; inflammatory disorder; allergy; developmental disorder; infection;  
KW hypothyroidism; cushing's syndrome; gene therapy; cytostatic; nootropic;  
KW anticonvulsant; neuroprotective; cerebroprotective; thyromimetic; gene;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
CDS 301..1074  
FT /\*tag= a  
FT /product= "Human NAAP-4 protein"  
FT sig\_peptide 301..420  
FT /\*tag= b  
FT mat\_peptide 421..1071  
FT /\*tag= c  
FT /product= "Human mature NAAP-4 protein"  
XX  
XX WO200299115-A2.  
XX  
XX 12-DEC-2002.  
XX  
XX 31-MAY-2002; 2002WO-US017050.  
XX  
XX 01-JUN-2001; 2001US-0295359P.  
XX 08-JUN-2001; 2001US-0296878P.  
XX 08-JUN-2001; 2001US-0297222P.  
XX 15-JUN-2001; 2001US-0298615P.  
XX 15-JUN-2001; 2001US-0298665P.  
XX 15-JUN-2001; 2001US-0298693P.  
XX 21-JUN-2001; 2001US-0300176P.  
XX 19-APR-2002; 2002US-0373891P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PA (YUEH/) YUE H.  
XX  
XX Yue H, Tang YT, Baughn MR, Becha SD, Warren BA, Walla NK,  
PI Lal PG, Lee EA, Hafalia AJA, Richardson TW, Griffin JA, Emerling BM;  
PI Ramkumar J, Yue H, Swarnakar A, Tran B, Li JX, Yao MG, Yang J;  
PI Ison CH, Forsythe IJ, Honchell CD, Arvizu CS, Elliott VS, Lu Y;  
PI Ding L, Luo W, Wang YE, Burford N, Borowsky ML, Nguyen DB;  
PI Chinn AM, Kable AE;  
XX

DR WPI; 2003-140626/13.  
DR P-PSDB; AAE33764.  
XX New human nucleic acid associated proteins (NAAP), useful for diagnosing,  
PT treating and preventing diseases or conditions associated with the  
PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or  
PT infections.  
XX  
PS Claim 89; Col 231; 257pp; English.  
XX The present invention relates to human nucleic acid associated proteins  
CC (NAAP) and polynucleotides encoding such proteins. NAAP sequences are  
CC useful in diagnosing, treating and preventing diseases or conditions  
CC associated with the decreased expression or overexpression of NAAP such  
CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.  
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,  
CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)  
CC disorders or infections. They are also used in gene therapy. The present  
CC sequence is human NAAP-4 cDNA  
XX  
SQ Sequence 1345 BP; 255 A; 425 C; 456 G; 209 T; 0 U; 0 Other;  
Query Match 79.0%; Score 15.8; DB 10; Length 1345;  
Best Local Similarity 89.5%; Pred. No. 5.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGTGTGCTGTGCTAGTCC 19  
DB 1134 CGTGTGCTGTGCTAGTCC 1116  
RESULT 28  
ABZ11458  
ID ABZ11458 standard; cDNA; 2304 BP.  
XX  
AC ABZ11458;  
XX  
DT 20-JAN-2003 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 340.  
XX  
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;  
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200270539-A2.  
XX  
PD 12-SEP-2002.  
XX  
PF 05-MAR-2002; 2002WO-US005095.  
XX  
PR 05-MAR-2001; 2001US-00799451.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX  
DR WPI; 2002-759812/82.  
DR P-PSDB; ABP69241.  
XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
PT or coagulation disorders.

XX  
PS Claim 1; SEQ ID NO 340; 1012pp + Sequence Listing; English.  
XX The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-  
CC ABZ12066) or their mature protein coding portion, active domain coding  
CC protein or complementary sequences. The polynucleotides are useful for  
CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell-  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
CC platelet or coagulation disorders, wound, burn, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2304 BP; 394 A; 800 C; 720 G; 400 T; 0 U; 0 Other;  
Query Match 79.0%; Score 15.8; DB 6; Length 2304;  
Best Local Similarity 89.5%; Pred. No. 5.8e+02;  
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGTGTGCTGTGCTAGTCC 19  
DB 651 CGTGTGCTGTGCTAGTCC 669  
RESULT 29  
ADM43976  
ID ADM43976 standard; cDNA; 2304 BP.  
XX  
AC ADM43976;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Novel human arginine-rich protein cDNA #340.  
XX  
KW ss; gene; human; arginine-rich protein; cancer; inflammation;  
KW genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2004053250-A1.  
XX  
PD 18-MAR-2004.  
XX  
PF 21-NOV-2002; 2002US-00302172.  
XX  
PR 05-MAR-2001; 2001US-00799451.  
PR 05-MAR-2002; 2002WO-US005095.  
PR 20-AUG-2002; 2002US-00225251.  
XX  
PA (TANG/) TANG Y T.  
PA (XUEA/) XUE A.  
PA (DRMA/) DRMANAC R T.  
XX  
PI Tang YT, Xue A, Drmanac RT;  
XX  
DR WPI; 2004-238579/22.  
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,  
PT useful for diagnosing and/or treating conditions associated with aberrant  
PT activity of the arginine-rich polypeptides, such as cancer and  
PT inflammation.  
XX  
PS Disclosure; SEQ ID NO 340; 51pp; English.  
XX The invention relates to an isolated polynucleotide. The methods and  
CC compositions of the present invention are useful for the diagnosis and/or

CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the arginine-rich protein-like polypeptides, such as  
 CC cancer and inflammation. They can also be used in forensics, gene  
 CC mapping, identification of mutations responsible for genetic disorders,  
 CC and in assessing biodiversity. The present sequence represents a novel  
 CC human arginine-rich protein cDNA.

XX SQ Sequence 2304 BP; 384 A; 800 C; 720 G; 400 T; 0 U; 0 Other;  
 Query Match 79.0%; Score 15.8; DB 12; Length 2304;  
 Best Local Similarity 89.5%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCC 19  
 DB 651 CGTGTCTCTGCGAGGCC 669

RESULT 30  
 AAI66465  
 ID AAI66465 standard; cDNA; 2434 BP.  
 XX AC AAI66465;  
 XX DT 04-DEC-2001 (first entry)  
 XX DE P53 tumour antigen protein 13 coding sequence.  
 XX KW P53 tumour antigen protein 13; cancer; haemopathy; HIV infection;  
 XX immunological disease; inflammation; gene therapy; ss.  
 XX OS Unidentified.  
 XX PN CN1301754-A.  
 XX PD 04-JUL-2001.  
 XX PF 29-DEC-1999; 99CN-00127233.  
 XX PR 29-DEC-1999; 99CN-00127233.  
 XX PA (UYFU-) UNIV FUDAN.  
 XX PI Mao Y, Xie Y;  
 XX DR WPI; 2001-558148/63.  
 XX P-PSDB; AAG78691.  
 XX PT New polypeptide-P53 tumor antigen protein 13 and a polynucleotide  
 XX encoding the polypeptide.  
 XX PS Claim 6; Page 23-24(Disclosure); 31pp; Chinese.  
 XX CC The present invention provides the protein and coding sequences of P53  
 CC tumour antigen protein 13. The sequences can be used in the treatment of  
 CC cancer, haemopathy, HIV infection, immunological diseases and  
 CC inflammation. The present sequence is the coding sequence of the  
 CC invention

XX SQ Sequence 2434 BP; 562 A; 728 C; 568 G; 576 T; 0 U; 0 Other;  
 Query Match 79.0%; Score 15.8; DB 4; Length 2434;  
 Best Local Similarity 89.5%; Pred. No. 5.8e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20  
 DB 783 GTGGGTCTGTGCTTGTCCC 801

RESULT 31  
 ABL09670  
 ID ABL09670 standard; cDNA; 3545 BP.

XX ABL09670;  
 AC 26-MAR-2002 (first entry)  
 DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 23492.  
 XX DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX OS Drosophila melanogaster.  
 XX PN WO200171042-A2.  
 XX PD 27-SEP-2001.  
 XX PF 23-MAR-2001; 2001WO-US009231.  
 XX PR 23-MAR-2000; 2000US-0191637P.  
 XX PR 11-JUL-2000; 2000US-00614150.  
 XX PA (PEKE ) PE CORP NY.  
 XX PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX DR WPI; 2001-656860/75.  
 XX P-PSDB; ABB65567.  
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX genes from Drosophila and for elucidating cell signalling and cell-cell  
 XX interactions.  
 XX PS Claim 1; SEQ ID NO 23492; 21pp + Sequence Listing; English.  
 XX CC The invention relates to an isolated nucleic acid detection reagent  
 XX capable of detecting 1000 or more genes from Drosophila. The invention is  
 XX useful in developmental biology and in elucidating cell signalling and  
 XX cell-cell interactions in higher eukaryotes for the development of  
 XX insecticides, therapeutics and pharmaceutical drugs. The invention  
 XX discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA  
 XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 XX ABB72072). The sequence data for this patent did not form part of the  
 XX printed specification, but was obtained in electronic format directly  
 XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 3545 BP; 966 A; 748 C; 742 G; 1089 T; 0 U; 0 Other;  
 Query Match 79.0%; Score 15.8; DB 4; Length 3545;  
 Best Local Similarity 89.5%; Pred. No. 6.1e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20  
 DB 2615 GTGTCTCTGCTAGTCCC 2633

RESULT 32  
 ADE55103  
 ID ADE55103 standard; DNA; 5315 BP.  
 XX AC ADE55103;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Rat gene AF059030, SEQ ID NO 908.  
 XX KW Rat; ds; gene; pain; neuronal tissue; gene therapy;  
 XX spinal segmental nerve injury; chronic constriction injury; CCI;  
 XX spared nerve injury; SNI; Chung.  
 XX OS Rattus norvegicus.  
 XX PN WO2003016475-A2.

XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX (GEO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX DR GENBANK; AF059030.  
XX New composition comprising two or more isolated polypeptides, useful for  
XX preparing a medicament for treating pain in an animal.  
XX Claim 1; Page; 1017pp; English.  
XX The invention discloses a composition comprising two or more isolated rat  
XX or human polynucleotides or a polynucleotide which represents a fragment,  
XX derivative or allelic variation of the nucleic acid sequence. Also  
XX claimed are a vector comprising the novel polynucleotide, a host cell  
XX comprising the vector, a method for identifying a nucleotide sequence  
XX which is differentially regulated in an animal subjected to pain and a  
XX kit to perform the method, an array, a method for identifying an agent  
XX that increases or decreases the expression of the polynucleotide sequence  
XX that is differentially expressed in neuronal tissue of a first animal  
XX subjected to pain, a method for identifying a compound which regulates  
XX the expression of a polynucleotide sequence which is differentially  
XX expressed in an animal subjected to pain, a method for identifying a  
XX compound that regulates the activity of one or more of the  
XX polynucleotides, a method for producing a pharmaceutical composition, a  
XX method for identifying a compound or small molecule that regulates the  
XX activity in an animal of one or more of the polypeptides given in the  
XX specification, a method for identifying a compound useful in treating  
XX pain and a pharmaceutical composition comprising the one or more  
XX polypeptides or their antibodies. The polynucleotide or the compound that  
XX modulates its activity is useful for preparing a medicament for treating  
XX pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction  
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
XX therapy). The sequence presented is a rat DNA (shown in Table 2 of the  
XX specification) which encodes one of the polypeptides of the invention  
XX which is differentially expressed during pain. Note: The sequence data  
XX for this patent did not form part of the printed specification, but was  
XX obtained in electronic form directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 5315 BP; 1349 A; 1324 C; 1288 G; 1354 T; 0 U; 0 Other;  
Query Match 79.0%; Score 15.8; DB 10; Length 5315;  
Best Local Similarity 89.5%; Pred. No. 6.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GTGTGTCGTGCTAGTCCC 20  
Db 3039 GTGGGTCTGTGCTAGTCTC 3057

RESULT 33  
ABL09674  
ID ABL09674 standard; cDNA; 5856 BP.  
XX ABL09674;  
XX ABL09674;  
XX 26-MAR-2002 (first entry)  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 23504.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX

KW pharmaceutical; gene; ss.  
XX Drosophila melanogaster.  
XX WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX P-PSDB; ABB65571.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions.  
XX Claim 1; SEQ ID NO 23504; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
XX ABB72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 5856 BP; 1527 A; 1429 C; 1445 G; 1455 T; 0 U; 0 Other;  
Query Match 79.0%; Score 15.8; DB 4; Length 5856;  
Best Local Similarity 89.5%; Pred. No. 6.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GTGTGTCGTGCTAGTCCC 20  
Db 421 GTGTGTCGTGCTAGTCCC 439

RESULT 34  
ABX09140  
ID ABX09140 standard; DNA; 82993 BP.  
XX ABX09140;  
XX ABX09140;  
XX 08-APR-2003 (first entry)  
XX Mycobacterium tuberculosis H37Rv BAC clone BAC-Rv221.  
XX Mycobacteriosis; survival; virulence; protective antigen; vaccine;  
XX mycobacterial disease; tuberculosis; leprosy; ds; cosmid.  
XX Mycobacterium tuberculosis.  
XX Mycobacterium tuberculosis.  
XX WO200274903-A2.  
XX 26-SEP-2002.  
XX 22-FEB-2002; 2002WO-IB001973.  
XX 22-FEB-2001; 2001US-0270123P.  
XX (INSP ) INST PASTEUR.  
XX

PI Cole S;  
XX  
DR WPI; 2002-759885/82.  
XX  
PT Identifying and selecting genes for survival or virulence of mycobacteria  
PT by a comparative genomic analysis of the sequences of Mycobacterium  
PT tuberculosis and M. leprae.  
XX  
XX Disclosure; Fig 4; 874pp; English.  
XX  
XX This invention relates to a novel method for identifying essential genes  
CC for survival or virulence of mycobacteria species. The method comprises  
CC aligning the genomic sequence of a first mycobacterium species on a  
CC genomic sequence of a second mycobacterium species and selecting a  
CC polynucleotide sequence that is highly conserved in both genomes with no  
CC counterparts in other bacterial genomic sequences and that corresponds to  
CC an essential gene for the survival or virulence of mycobacterium species.  
CC The method of the invention is useful for detecting M. tuberculosis or M.  
CC leprae infection. The method reduces the number of potential new targets  
CC and protective antigens for new drugs and vaccine compositions to treat  
CC and prevent mycobacterial diseases, particularly tuberculosis and  
CC leprosy. The present sequence represents a Mycobacterial cosmid DNA  
CC sequence used in the method of the invention  
XX

SQ Sequence 82993 BP; 14881 A; 2628 C; 26719 G; 14765 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 6; Length 82993;  
Best Local Similarity 89.5%; Pred. No. 9.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCC 20  
DB 29325 GTGTCTCTGTCTGTCG 29343

RESULT 35  
AAI99682\_21  
Continuation (22 of 45) of AAI99682 from base 2100001 (Mycobacterium tuberculosis strain  
WP Sequence split into 45 fragments LOCUS AAI99682 Accession AAI99682

WP	Fragment Name	Begin	End
WP	AAI99682_00	1	110000
WP	AAI99682_01	100001	210000
WP	AAI99682_02	200001	310000
WP	AAI99682_03	300001	410000
WP	AAI99682_04	400001	510000
WP	AAI99682_05	500001	610000
WP	AAI99682_06	600001	710000
WP	AAI99682_07	700001	810000
WP	AAI99682_08	800001	910000
WP	AAI99682_09	900001	1010000
WP	AAI99682_10	1000001	1110000
WP	AAI99682_11	1100001	1210000
WP	AAI99682_12	1200001	1310000
WP	AAI99682_13	1300001	1410000
WP	AAI99682_14	1400001	1510000
WP	AAI99682_15	1500001	1610000
WP	AAI99682_16	1600001	1710000
WP	AAI99682_17	1700001	1810000
WP	AAI99682_18	1800001	1910000
WP	AAI99682_19	1900001	2010000
WP	AAI99682_20	2000001	2110000
WP	AAI99682_21	2100001	2210000
WP	AAI99682_22	2200001	2310000
WP	AAI99682_23	2300001	2410000
WP	AAI99682_24	2400001	2510000
WP	AAI99682_25	2500001	2610000
WP	AAI99682_26	2600001	2710000
WP	AAI99682_27	2700001	2810000
WP	AAI99682_28	2800001	2910000
WP	AAI99682_29	2900001	3010000
WP	AAI99682_30	3000001	3110000
WP	AAI99682_31	3100001	3210000
WP	AAI99682_32	3200001	3310000

WP	AAI99682_33	3300001	3410000
WP	AAI99682_34	3400001	3510000
WP	AAI99682_35	3500001	3610000
WP	AAI99682_36	3600001	3710000
WP	AAI99682_37	3700001	3810000
WP	AAI99682_38	3800001	3910000
WP	AAI99682_39	3900001	4010000
WP	AAI99682_40	4000001	4110000
WP	AAI99682_41	4100001	4210000
WP	AAI99682_42	4200001	4310000
WP	AAI99682_43	4300001	4410000
WP	AAI99682_44	4400001	4411529

Query Match 79.0%; Score 15.8; DB 4; Length 110000;  
Best Local Similarity 89.5%; Pred. No. 9.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCC 20  
DB 44936 GTGTCTCTGTCTGTCG 44954

RESULT 36

AAI99683\_21

Continuation (22 of 44) of AAI99683 from base 2100001 (Mycobacterium tuberculosis strain  
WP Sequence split into 44 fragments LOCUS AAI99683 Accession AAI99683

WP	Fragment Name	Begin	End
WP	AAI99683_00	1	110000
WP	AAI99683_01	100001	210000
WP	AAI99683_02	200001	310000
WP	AAI99683_03	300001	410000
WP	AAI99683_04	400001	510000
WP	AAI99683_05	500001	610000
WP	AAI99683_06	600001	710000
WP	AAI99683_07	700001	810000
WP	AAI99683_08	800001	910000
WP	AAI99683_09	900001	1010000
WP	AAI99683_10	1000001	1110000
WP	AAI99683_11	1100001	1210000
WP	AAI99683_12	1200001	1310000
WP	AAI99683_13	1300001	1410000
WP	AAI99683_14	1400001	1510000
WP	AAI99683_15	1500001	1610000
WP	AAI99683_16	1600001	1710000
WP	AAI99683_17	1700001	1810000
WP	AAI99683_18	1800001	1910000
WP	AAI99683_19	1900001	2010000
WP	AAI99683_20	2000001	2110000
WP	AAI99683_21	2100001	2210000
WP	AAI99683_22	2200001	2310000
WP	AAI99683_23	2300001	2410000
WP	AAI99683_24	2400001	2510000
WP	AAI99683_25	2500001	2610000
WP	AAI99683_26	2600001	2710000
WP	AAI99683_27	2700001	2810000
WP	AAI99683_28	2800001	2910000
WP	AAI99683_29	2900001	3010000
WP	AAI99683_30	3000001	3110000
WP	AAI99683_31	3100001	3210000
WP	AAI99683_32	3200001	3310000
WP	AAI99683_33	3300001	3410000
WP	AAI99683_34	3400001	3510000
WP	AAI99683_35	3500001	3610000
WP	AAI99683_36	3600001	3710000
WP	AAI99683_37	3700001	3810000
WP	AAI99683_38	3800001	3910000
WP	AAI99683_39	3900001	4010000
WP	AAI99683_40	4000001	4110000
WP	AAI99683_41	4100001	4210000
WP	AAI99683_42	4200001	4310000
WP	AAI99683_43	4300001	4403765

Query Match

79.0%; Score 15.8; DB 4; Length 110000;

```
Best Local Similarity 89.5%; Pred. NO. 9.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTCTGTCTAGTCCC 20
DB 42152 GTGTGCTCTGTCTGTCTGCG 42170

RESULT 37
ADL13471
ID ADL13471 standard; DNA; 214520 BP.
XX AC ADL13471;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #3.
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
XX osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.
XX Disclosure; SEQ ID NO 3; 297pp; English.
XX The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
XX Sequence 214520 BP; 57899 A; 52039 C; 51001 G; 53577 T; 0 U; 4 Other;

Query Match 79.0%; Score 15.8; DB 10; Length 214520;
Best Local Similarity 89.5%; Pred. NO. 1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCC 19
DB 213527 CGTGTCTCTGTCTAGTCC 213545

RESULT 38
```

```
ADP43517
ID ADP43517 standard; DNA; 347001 BP.
XX AC ADP43517;
XX DT 09-SEP-2004 (first entry)
XX DE Human MAD1-like 1 DNA #7.
XX ds; gene; human.
XX KW Homo sapiens.
XX OS US2004115650-A1.
XX PN 17-JUN-2004.
XX PD 12-DEC-2002; 2002US-00319908.
XX PF 12-DEC-2002; 2002US-00319908.
XX PR (ISIS-) ISIS PHARM INC.
XX PA Dobie KW, Jain R;
XX PI WPI; 2004-449387/42.
XX DR New oligonucleotide compound that inhibits expression of MAD1-like 1,
PT useful for preparing a composition for treating hyperproliferative
PT disorder, e.g., cancer.
XX Example 15; SEQ ID NO 16; 206pp; English.
XX The invention relates to a new compound targeted to a nucleic acid
CC encoding MAD1-like 1 which specifically hybridises with the nucleic acid
CC encoding MAD1-like 1 and inhibits expression of MAD1-like 1. The
CC oligonucleotide compound is useful for preparing a composition for
CC treating hyperproliferative disorder, e.g. cancer. The present sequence
CC represents human MAD1-like 1 DNA.
XX Sequence 347001 BP; 64746 A; 89476 C; 97054 G; 85725 T; 0 U; 10000 Other;

Query Match 79.0%; Score 15.8; DB 12; Length 347001;
Best Local Similarity 89.5%; Pred. NO. 1.1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTCTGTCTAGTCCC 20
DB 237103 GTGTGCTCTGTCTGTCTCC 237121

RESULT 39
ADES59507/c
ID ADE59507 standard; DNA; 419 BP.
XX AC ADE59507;
XX DT 29-JAN-2004 (first entry)
XX DE Rat gene AA891308, SEQ ID NO 5402.
XX Rat; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX
```

```

PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GHEO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; AA891308.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the
CC specification) which is differentially expressed during pain. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 419 BP; 113 A; 107 C; 87 G; 112 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 10; Length 419;
Best Local Similarity 94.1%; Pred. No. 7.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGCTGCTGCTAGTC 18
DB 175 GTGTGCTGCTGCTGTC 159

RESULT 40
ADJ10940/C
ID ADJ10940 standard; DNA; 448 BP.
XX
AC ADJ10940;
XX
DT 17-JUN-2004 (first entry)
XX
DE Recombinant tomato DNA to generate disease resistant plants SeqID 336.
XX
XX Genetically engineered; plant; Avr-Pto; avirulent Pto;
KW bacterial speck disease; tomato; transgenic; biotic; abiotic stress; ds;
KW -Prf.
XX
XX Lycopersicon esculentum.
OS
XX
XX US2004006787-A1.
PN

```

---

```

XX 08-JAN-2004.
PD
XX
PF 14-JAN-2003; 2003US-00341961.
XX
XX 14-JAN-2002; 2002US-0348792P.
PR
PR 20-JUN-2002; 2002US-0390249P.
XX
XX (MART/) MARTIN G B.
PA (MYSO/) MYSOORE K K.
PA (CRAS/) CRASTA O R.
PA (FOLK/) FOLKERTS O.
PA (SWIR/) SWIRSKY P.
XX
PI Martin GB, Mysore KK, Crasta OR, Folkerts O, Swirsky P;
XX
XX WPI; 2004-081759/08.
XX
XX New genetically engineered plant comprising a recombinant polynucleotide
PT showing expression associated with Avr-Pto mediated defense response,
PT useful in increasing resistance of plant against bacterial speck disease.
XX
XX Claim 5; SEQ ID NO 336; 176pp; English.
XX
XX This invention relates to a method of generating novel genetically
CC engineered plants. Specifically, it refers to compositions useful for
CC transforming plants with a recombinant polynucleotide or plant gene that
CC shows a specific pattern of expression associated with the Avr-Pto
CC mediated defence response. Avr-Pto is described as the avirulent Pto gene
CC found in Pseudomonas syringae pathovar tomato strain [strain Tl(A)],
CC which mediates bacterial speck disease in tomato plants. The present
CC invention describes identifying genes that are up- or down- regulated in
CC the defence response and that are involved in the interaction of Avr-Pto
CC with Pto and/ or Prf. As such, these novel recombinant polynucleotides
CC can be used to generate transgenic plants that are resistant to bacterial
CC speck disease and furthermore exhibit increased resistance against biotic
CC and abiotic stresses. This polynucleotide sequence is a recombinant
CC tomato polynucleotide used to generate genetically engineered, disease
CC resistant plants of the invention.
XX
SQ Sequence 448 BP; 131 A; 102 C; 99 G; 115 T; 0 U; 1 Other;

Query Match 77.0%; Score 15.4; DB 12; Length 448;
Best Local Similarity 94.1%; Pred. No. 7.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTGTGCTGCTGCTAGT 17
DB 116 CGTGTGCTGCTGCTAGT 100

RESULT 41
ABZ12515
ID ABZ12515 standard; DNA; 1644 BP.
XX
AC ABZ12515;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 320.
XX
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
KW Arabidopsis thaliana.
OS
XX
XX WO200216655-A2.
PN
XX
XX 28-FEB-2002.
PD
XX
XX 24-AUG-2001; 2001WO-US026685.
PF
XX
XX 24-AUG-2000; 2000US-0227866P.
PR
PR 26-JAN-2001; 2001US-0264647P.
PR

```

PR 22-JUN-2001; 2001US-0300111P.  
 PA (SCRI ) SCRIPPS RES INST.  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX Harper JF, Kreps J, Wang X, Zhu T;  
 XX WPI; 2002-304127/34.  
 DR  
 XX  
 PT Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.  
 XX  
 PS Claim 144; SEQ ID NO 320; 577pp + Sequence Listing; English.  
 XX  
 CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 SQ Sequence 1644 BP; 361 A; 358 C; 463 G; 462 T; 0 U; 0 Other;  
 Query Match 77.0%; Score 15.4; DB 6; Length 1644;  
 Best Local Similarity 94.1%; Pred. No. 8.8e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GTGTCGTGCTAGTCCC 20  
 DB 1504 GTGATGTGCTAGTCCC 1520  
 RESULT 42  
 ABL41879/c  
 ID ABL41879 standard; DNA; 1690 BP.  
 XX  
 AC ABL41879;  
 XX  
 DT 11-JUN-2002 (first entry)  
 XX  
 DE Nucleotide sequence of a linear polyol transporter.  
 XX  
 KW Linear polyol; mannitol; polyol transporter; carbon source; plant;  
 KW pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;  
 KW inositol; ribitol; xylitol; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200204647-A1.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 22-JUN-2001; 2001WO-FR001979.  
 XX  
 PR 11-JUL-2000; 2000FR-00009032.  
 XX  
 PA (CNRS ) CENT NAT RECH SCI.  
 XX  
 PI Lemoine RRP, Noiraud NEJ;  
 XX  
 DR WPI; 2002-154933/20.  
 XX  
 PT New polyol transporter protein from plants, for selecting transformed  
 PT cells and for imparting pathogen and salt-stress resistance to plants.  
 XX  
 PS Disclosure; Page 45; 66pp; French.  
 XX

CC The present sequence encodes a polypeptide which is a transporter of a  
 CC linear polyol. DNA encoding linear polyols is used to produce a selection  
 CC system for transformed cells, based on the use of polyols as the only  
 CC carbon source. It is also used to produce transgenic plants with  
 CC increased resistance to pathogens and salt stress. The use of a linear  
 CC polyol for selection eliminates the need for toxic selection reagents  
 CC such as antibiotics. Also, the linear polyol is not essential for the  
 CC plant, once selection has been made. The linear polyol has a main chain  
 CC of 5-8, preferably 6, carbon atoms and is selected from sorbitol,  
 CC dulcitol, galactitol, (myo-)inositol, ribitol, or xylitol  
 XX  
 SQ Sequence 1690 BP; 479 A; 480 C; 365 G; 366 T; 0 U; 0 Other;  
 Query Match 77.0%; Score 15.4; DB 6; Length 1690;  
 Best Local Similarity 94.1%; Pred. No. 8.9e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GTGTCGTGCTAGTCCC 20  
 DB 141 GTGATGTGCTAGTCCC 125  
 RESULT 43  
 AAX20533  
 ID AAX20533 standard; DNA; 5819 BP.  
 XX  
 AC AAX20533;  
 XX  
 DT 05-MAY-1999 (first entry)  
 XX  
 DE Polynucleotide sequence from the genome of Treponema pallidum.  
 XX  
 KW Treponema pallidum infection; syphilis; Borrelia infection; animal;  
 KW enzyme production; ds.  
 XX  
 OS Treponema pallidum.  
 XX  
 PN WO9859034-A2.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 23-JUN-1998; 98WO-US013041.  
 XX  
 PR 24-JUN-1997; 97US-0050667P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Fraser CM;  
 XX  
 DR WPI; 1999-081273/07.  
 XX  
 PT New isolated Treponema pallidum nucleic acids - used to develop products  
 PT for the detection, diagnosis, characterisation, prevention and therapy of  
 PT T. pallidum infections, particularly syphilis.  
 XX  
 PS Claim 1; Page 371-374; 1150pp; English.  
 XX  
 CC AAX20500-21243 represent polynucleotide sequences from the genome of  
 CC Treponema pallidum. The sequences can be used for detection, diagnosis,  
 CC characterisation, prevention and therapy for T. pallidum infections,  
 CC particularly syphilis. They can also be used for detecting diseases  
 CC related to Borrelia infections in animals, and for the production of  
 CC biosynthetic products such as enzymes  
 XX  
 SQ Sequence 5819 BP; 1257 A; 1273 C; 1734 G; 1549 T; 0 U; 6 Other;  
 Query Match 77.0%; Score 15.4; DB 2; Length 5819;  
 Best Local Similarity 88.9%; Pred. No. 1e+03;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GTGTCGTGCTAGTCCC 19  
 DB 4816 GTGTCGTGCTAGTCCC 4833



## RESULT 44

ADA02531/c  
ID ADA02531 standard; DNA; 21981 BP.

XX AC ADA02531;

XX DT 06-NOV-2003 (first entry)

XX DE Mouse Sox4 carcinoma associated gene, SEQ ID NO:1049.

XX KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;  
XX KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
XX KW gene; ds.

XX OS Mus sp.

XX PN WO2003057146-A2.

XX XX 17-JUL-2003.

XX PF 26-DEC-2002; 2002WO-US041414.

XX PR 26-DEC-2001; 2001US-00035832.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX DR WPI; 2003-587068/55.

XX PT New recombinant nucleic acid encoding carcinoma associated protein,  
XX PT useful for preparing compositions for treating carcinomas.

XX PS Claim 1; SEQ ID NO 1049; 245pp; English.

XX CC The invention relates to recombinant carcinoma associated (CA) nucleic  
XX CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
XX CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
XX CC invention also encompasses expression vectors and host cells comprising a  
XX CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
XX CC binds to the protein, and a biochip comprising CA nucleic acid or  
XX CC fragments thereof. The sequences of the invention were identified using  
XX CC oncogenic retroviruses, which insert into the genome of the host organism  
XX CC at random. Many of these do not carry transduced host oncogenes or  
XX CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
XX CC direct consequence of the effects of proviral integration into host  
XX CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
XX CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
XX CC leukaemia) or a propensity to carcinoma by determination of the sequence  
XX CC of a CA gene, or by determination of CA gene expression in particular  
XX CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
XX CC therapeutic agents and in screening and evaluating drug candidates. The  
XX CC present sequence represents a specifically claimed murine CA nucleic acid  
XX CC sequence of the invention. Note: The complete sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained  
XX CC in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 21981 BP; 5568 A; 4945 C; 5268 G; 5590 T; 0 U; 610 Other;

Query Match 77.0%; Score 15.4; DB 9; Length 21981;

Best Local Similarity 94.1%; Pred. No. 1.2e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTACT 17

DB 8879 CGTGTCTGTGCTACT 8863

## RESULT 45

ADB72269/c

ID ADB72269 standard; DNA; 21981 BP.

XX AC ADB72269;

XX DT 04-DEC-2003 (first entry)

XX DE Mouse Sox4 gene.

XX KW mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;  
XX KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.

XX OS Mus sp.

XX PN WO2003008583-A2.

XX PD 30-JAN-2003.

XX PF 26-DEC-2001; 2001WO-US051291.

XX PR 02-MAR-2001; 2001US-00798586.

XX PR 23-OCT-2001; 2001US-00004113.

XX PR 08-NOV-2001; 2001US-00052482.

XX PR 30-NOV-2001; 2001US-00997722.

XX PR 20-DEC-2001; 2001US-00034650.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW, Engelhard EK;

XX DR WPI; 2003-239337/23.

XX PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,  
XX PT cancers, neoplasm, adenocarcinoma, or sarcomas.

XX PS Claim 1; SEQ ID NO 97; 2304pp; English.

XX CC The invention relates to a novel recombinant nucleic acid comprising a  
XX CC nucleotide sequence selected from any of the 660 sequences fully defined  
XX CC in the specification. A polynucleotide of the invention has cytostatic  
XX CC activity, and may have a use in gene therapy, or in a vaccine. The  
XX CC recombinant nucleic acids and polypeptides are useful for treating  
XX CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and  
XX CC sarcomas. The present sequence represents a mouse gene of the invention.

XX SQ Sequence 21981 BP; 5568 A; 4945 C; 5268 G; 5590 T; 0 U; 610 Other;

Query Match 77.0%; Score 15.4; DB 10; Length 21981;

Best Local Similarity 94.1%; Pred. No. 1.2e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTACT 17

DB 8879 CGTGTCTGTGCTACT 8863

## RESULT 46

ADE95779/c

ID ADE95779 standard; DNA; 21981 BP.

XX AC ADE95779;

XX DT 12-FEB-2004 (first entry)

XX DE Mouse Sox4 gene genomic DNA sequence.

XX KW cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;  
XX KW lymphoma; breast cancer; prostate cancer; leukaemia; ds; mouse; murine;  
XX KW Sox4.

XX OS Mus sp.

XX PN WO2003039484-A2.

PD 15-MAY-2003.  
XX  
PF 08-NOV-2002; 2002WO-US036071.  
XX  
XX 08-NOV-2001; 2001US-00052482.  
XX  
PA (SAGR-) SAGRES DISCOVERY.  
XX  
XX Morris DW, Engelhard EK;  
PI WPI; 2003-441462/41.  
XX  
DR New carcinoma associated nucleic acids and proteins, useful for screening  
PT drug candidates, or for diagnosing and treating carcinomas, e.g.  
PT lymphoma, breast cancer, prostate cancer or leukemia.  
XX  
XX Claim 1; SEQ ID NO 37; 793bp; English.  
XX  
XX This invention relates to novel recombinant nucleic acids for use in  
CC diagnosis and treatment of cancer, especially carcinomas, as well as the  
CC use of compositions in screening methods. The compositions of the  
CC invention may have cytostatic activity whilst the disclosed sequences may  
CC be useful for gene therapy. The carcinoma associated nucleic acids and  
CC proteins are useful for diagnosing and treating carcinomas, for example  
CC lymphoma, breast cancer, prostate cancer or leukaemia, or for screening  
CC drug candidates or bioactive agents capable of binding to, or modulating  
CC the activity of, a carcinoma associated protein. The present sequence is  
CC the genomic DNA sequence of the mouse Sox4 gene which is a carcinoma  
CC associated gene of the invention.  
XX  
XX Sequence 21981 BP; 5568 A; 4945 C; 5268 G; 5590 T; 0 U; 610 Other;  
SQ  
Query Match 77.0%; Score 15.4; DB 10; Length 21981;  
Best Local Similarity 94.1%; Pred. No. 1.2e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGTGTCTGTGTCTACT 17  
DB 8879 CGTGTCTGTGTCTACT 8863  
RESULT 47  
ABD32841  
ID ABD32841 standard; DNA; 227448 BP.  
XX  
XX ABD32841;  
AC  
XX 18-NOV-2004 (first entry)  
DT  
XX Mouse cancer-associated genomic DNA MD17-014.  
DE  
XX Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;  
KW leukaemia; lymphoma; CAP.  
XX  
XX Mus musculus.  
OS  
XX WO2004074320-A2.  
PN  
XX  
XX 02-SEP-2004.  
PD  
XX  
XX 17-FEB-2004; 2004WO-US004730.  
XX  
XX 14-FEB-2003; 2003US-00367094.  
PR  
XX 14-MAR-2003; 2003US-00388838.  
PR  
XX 15-APR-2003; 2003US-00417375.  
PR  
XX 13-JUN-2003; 2003US-00461862.  
PR  
XX 15-SEP-2003; 2003US-00663431.  
PR  
XX 15-DEC-2003; 2003US-00737318.  
XX  
XX (SAGR-) SAGRES DISCOVERY INC.  
PA  
XX Morris DW, Morris DW, Malandro MS;  
PI  
XX

DR WPI; 2004-652914/63.  
XX  
XX New isolated cancer-associated polynucleotides and polypeptides useful  
PT for diagnosing, preventing or treating cancers, especially lymphoma and  
PT leukemia, or in screening for agents that modulate cancer.  
XX  
XX disclosure; seqid 503; 310pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising at least 10  
CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
CC in the specification, or its complement. The nucleic acids encode cancer-  
CC associated proteins. Also included are an expression vector comprising  
CC the isolated nucleic acid cited above, a host cell comprising the above  
CC recombinant nucleic acid or expression vector, a microarray for detecting  
CC a cancer-associated (CA) nucleic acid comprising at least one probe  
CC comprising at least 10 contiguous nucleotides of any of the above-  
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
CC an open reading frame of a CA sequence selected from any of the 95  
CC polynucleotide sequences as mentioned in the specification, or its  
CC complement), an isolated antibody, (or its antigen binding fragment) that  
CC binds to the above polypeptide, a hybridoma that produces the above  
CC monoclonal antibody, a pharmaceutical composition comprising the above  
CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
CC cells (comprising the antibody cited above, methods for diagnosing cancer  
CC or for detecting the presence or absence of cancer cells in an  
CC individual, a method for inhibiting growth of cancer cells in an  
CC individual, a method for delivering a therapeutic agent to cancer cells  
CC in an individual, an electronic library comprising the above  
CC polynucleotide or polypeptide (or their fragments), methods of screening  
CC for anticancer activity or for a bioactive agent capable of modulating  
CC the activity of a CA protein (CAP), methods for detecting cancer  
CC associated with expression of a polypeptide in a test cell sample, a  
CC method for treating cancers and a method for inhibiting the expression of  
CC CA gene in a cell. The composition and methods are useful for detecting,  
CC diagnosing, preventing and treating cancers, especially lymphoma and  
CC leukaemia. These may also be used in screening for agents that modulate  
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 227448 BP; 54780 A; 54263 C; 56236 G; 60311 T; 0 U; 1858 Other;  
SQ  
Query Match 77.0%; Score 15.4; DB 13; Length 227448;  
Best Local Similarity 94.1%; Pred. No. 1.6e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 TGTGTCTGTGTCTAGTCC 19  
DB 198747 TGTGTCTGTGTAGTCC 198763  
RESULT 48  
ABQ74179  
ID ABQ74179 standard; DNA; 229354 BP.  
XX  
XX ABQ74179;  
AC  
XX 13-OCT-2002 (first entry)  
DT  
XX Human cytomegalovirus strain AD169 genomic sequence SEQ ID NO:64.  
DE  
XX Human cytomegalovirus; HCMV; CMV; Yeast artificial chromosome; YAC;  
KW vaccine; immunisation; infection; antiinflammatory; immunosuppressive;  
KW virucide; hepatotropic; congenital infection; jaundice; pneumonitis;  
KW respiratory distress; convulsive seizure; mental retardation; hepatitis;  
KW neurologic disability; asymptomatic adult infection; mononucleosis;  
KW retinitis; acquired immune deficiency syndrome; AIDS; genome; ds.  
XX  
XX Human cytomegalovirus.  
OS  
XX  
XX Key Location/Qualifiers  
FH promoter 80996. .81441  
FT

```
FT FT /*tag= a
FT FT /note= "UL54 promoter"
FT FT 128318.128750
FT FT /*tag= b
FT FT /note= "UL86 promoter"
XX XX
PN WO200257437-A2.
XX XX
XX 25-JUL-2002.
XX XX
XX 30-OCT-2001; 2001WO-US047943.
XX XX
XX 03-NOV-2000; 2000US-00705400.
XX XX
XX (SCRI ) SCRIPPS RES INST.
XX XX
XX Ghazal P, Huang H;
XX XX
XX WPI; 2002-590735/63.
XX XX
XX Cytomegalovirus-yeast artificial chromosome (CMV-YAC) DNA molecule useful
PT as a vaccine for treating CMV infections e.g. jaundice, respiratory
PT distress, convulsive seizures, mononucleosis, hepatitis, pneumonitis and
PT retinitis.
XX XX
PS Claim 5; Page 64-131; 132pp; English.
XX XX
CC The present invention describes an isolated recombinant DNA molecule
CC comprising a yeast artificial chromosome (YAC) including at least a
CC portion of a human cytomegalovirus (HCMV) genome. Also described is a
CC composition suitable for use as a vaccine comprising the isolated
CC recombinant DNA molecule and an excipient. The YAC has antiinflammatory,
CC immunosuppressive, virucide, and hepatotropic activities, and can be used
CC in vaccine production. The YAC vector and the recombinant DNA molecule
CC are useful as vaccines for treating infections caused by CMV, e.g.
CC congenital infections such as jaundice, respiratory distress and
CC convulsive seizures, which may result in mental retardation, neurologic
CC disability or death; or asymptomatic adult infections such as
CC mononucleosis, hepatitis, pneumonitis and retinitis. The vaccine can also
CC be used for immuno-compromised patients suffering from acquired immune
CC deficiency syndrome (AIDS), or undergoing chemotherapy or tissue
CC rejection therapy after organ transplantation. The vector is also useful
CC as a tool for studying the replication of CMV using cell-line models and
CC the function of essential cis acting genes or nucleic acid sequences. The
CC YAC vector provides an ease and rapidity in introducing single or
CC multiple mutations into a CMV DNA sequence. The capacity of the vectors
CC to generate multiple mutations exceeds the capacity of other vectors,
CC e.g. bacterial artificial chromosome. The present sequence represents
CC human cytomegalovirus strain AD169 genomic sequence (also see Genbank
CC X17403.1), given in the present invention
XX XX
SQ Sequence 229354 BP; 49475 A; 64911 C; 66192 G; 48776 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 6; Length 229354;
Best Local Similarity 94.1%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GGTGTCTGTGCTAGTC 18
DB 225853 GGTGTCTGTGCTAGTC 225869
RESULT 49
ACI35409/c
ID ACI35409 standard; DNA; 25 BP.
XX XX
XX ACI35409;
XX XX
XX 13-OCT-2003 (first entry)
XX XX
XX Human microarray DNA oligonucleotide SEQ ID NO 35400.
XX XX
XX EST; ss; probe; expressed sequence tag; microarray; gene expression;
```

```
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX Homo sapiens.
OS
XX US2003104410-A1.
XX 05-JUN-2003.
XX 15-MAR-2002; 2002US-00098263.
XX 16-MAR-2001; 2001US-0276759P.
XX (AFFY-) AFFYMETRIX INC.
XX Mittmann MP;
XX WPI; 2003-567953/53.
XX
XX New array of nucleic acid probes, useful for in situ hybridization, in
XX Southern, Northern or dot-blot hybridization to identify or detect the
XX sequence or specific mutations of any gene.
XX Claim 1; SEQ ID NO 35400; 9pp; English.
XX
XX The invention discloses a microarray comprising a plurality of nucleic
XX acid probes including one of 2,018,500 fully defined sequences, or its
XX perfect match, perfect mismatch, antisense match or antisense mismatch.
XX Also disclosed is a method of gene expression analysis. The array is used
XX in monitoring gene expression levels by hybridisation to a DNA library,
XX in analysis of genetic variation or in hybridisation of tag-labelled
XX compounds. The nucleic acid probes are specifically designed for analysis
XX of at least one target sequence. The method of analysis comprises
XX hybridising at least one or more nucleic acids to at least two or more
XX nucleic acid probes and detecting the hybridisation. The nucleic acid
XX probes are attached to a solid support. The analysis comprises monitoring
XX gene expression levels, identifying biallelic markers or polymorphisms,
XX or family members of a gene and a cross-species comparison. Each of the
XX nucleic acids further comprises a tag sequence. The array of nucleic acid
XX probes is useful in in situ hybridisation, in Southern, Northern or dot-
XX blot hybridisation to identify or detect the sequence or specific
XX mutations of any gene, in mapping the 5' termini of mRNA molecules by
XX primer extensions or in screening cDNA or genomic libraries or subclones
XX for additional subclones containing segments of DNA that have been
XX isolated and previously sequenced. The sequence presented is one of the
XX nucleic acid probes incorporated in the microarray. Note: The sequence
XX data for this patent can also be obtained in electronic format directly
XX from USPTO at seqdata.uspto.gov/sequence.html
XX
SQ Sequence 25 BP; 9 A; 5 C; 7 G; 4 T; 0 U; 0 Other;
Query Match 76.0%; Score 15.2; DB 9; Length 25;
Best Local Similarity 85.0%; Pred. No. 6.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGCTAGTCCC 20
DB 20 CATGTTCTGTCTAGTCCC 1
RESULT 50
ABN39874/c
ID ABN39874 standard; DNA; 60 BP.
XX XX
XX AC ABN39874;
XX XX
XX 15-JUL-2002 (first entry)
XX XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:12622.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
```

OS Homo sapiens.  
XX  
FN WO200210449-A2.  
XX  
PD  
XX 07-FEB-2002.  
XX  
PF 20-JUL-2001; 2001WO-IB001903.  
XX  
PR 28-JUL-2000; 2000US-0221607P.  
PR 02-MAY-2001; 2001US-0287724P.  
XX  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
PI WPI; 2002-257383/30.  
XX  
XX  
XX New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of a  
PT genome, useful for detecting tissue-, pathology-, and developmental-  
PT specific genes.  
XX  
XX Example 1; SEQ ID NO 12622; 47pp; English.  
XX  
XX The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
CC )transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises several  
CC oligonucleotides, each capable of hybridising selectively to a set of  
CC messenger RNAs transcribed from a given transcription unit of the genome,  
CC which encodes one or more messenger RNA splice variants. The  
CC oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterising the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a particular  
CC biological or pathological state, and so allowing the detection of tissue  
CC - and pathology-specific genes such as those genes only expressed in  
CC specific tissue under a specific pathological condition; to detect  
CC developmental specific genes; and to detect RNA transcripts and splice  
CC variants of a transcriptome of a patient suffering from a particular  
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from  
CC rats, humans and mice, which are used in the exemplification of the  
CC present invention. N.B. The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 60 BP; 18 A; 9 C; 20 G; 13 T; 0 U; 0 Other;

Query Match 76.0%; Score 15.2; DB 6; Length 60;  
Best Local Similarity 85.0%; Pred. No. 7.2e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCCC 20  
Db 53 CCTGTCTCTGCTAGTCCC 34

Search completed: March 15, 2005, 18:32:45  
Job time : 458 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 15:23:29 ; Search time 1672 Seconds  
(without alignments)  
579.608 Million cell updates/sec

Title: US-09-980-953-256  
Perfect score: 20  
Sequence: 1 cgtgtctgtctgtatgccc 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_scs.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	124	6	AR146422 Sequence
C 2	20	100.0	124	6	AR381504 Sequence
C 3	20	100.0	430	6	AX884562 Sequence
C 4	20	100.0	430	6	BD024172 Sequence
C 5	20	100.0	478	9	HSB7281
C 6	20	100.0	1112	6	AR534819 Sequence
C 7	20	100.0	1112	9	HUMB72A
C 8	20	100.0	1120	6	AR030780 Sequence
C 9	20	100.0	1120	6	AR112747 Sequence
C 10	20	100.0	1120	6	AR146413 Sequence
C 11	20	100.0	1120	6	BD272169 Use of so
C 12	20	100.0	1120	6	AR196804 Sequence
C 13	20	100.0	1120	6	AR287728 Sequence
C 14	20	100.0	1120	6	AR374141 Sequence
C 15	20	100.0	1120	6	AR381495 Sequence
C 16	20	100.0	1120	6	AR527666 Sequence
C 17	20	100.0	1120	6	AX047043 Sequence
C 18	20	100.0	2781	9	BC040261 Homo sapi
C 19	20	100.0	164161	9	AC068630 Homo sapi

20	100.0	178804	2	AC016918	AC016918 Homo sapi
21	17.4	2710	8	AY080604	AY080604 Arabidops
22	17.4	3168	6	CQ803940	CQ803940 Sequence
23	17.4	3518	8	AJ583496	AJ583496 Arabidops
24	17.4	91924	8	ATAC009606	ATAC009606 Arabidops
25	17.4	14858	2	AC140972	AC140972 Felis cat
26	17.4	146492	8	AC136223	AC136223 Oryza sat
27	17.4	161516	2	CR391906	CR391906 Danio rer
28	17.4	87.0	2	AC129160	AC129160 Rattus no
29	17.4	87.0	2	AC094647	AC094647 Rattus no
30	17.4	87.0	2	AC129648	AC129648 Rattus no
31	17.4	85.0	2	AC113396	AC113396 Mus muscu
32	17.4	85.0	2	LMFLCHR15_6	Continuation (7 of
33	17.4	85.0	2	CR382128_04	Continuation (5 of
34	16.8	84.0	5	BC076559	BC076559 Danio rer
35	16.8	84.0	9	AF220172S2	AF220172 Homo sapi
36	16.8	84.0	2	AC118055	AC118055 Homo sapi
37	16.8	84.0	2	AC118055	AC118055 Homo sapi
38	16.8	84.0	3	AC004252	AC004252 Drosophill
39	16.8	84.0	2	AC017870	AC017870 Drosophill
40	16.8	84.0	2	AC151836_3	Continuation (4 of
41	16.8	84.0	2	AC151419	AC151419 Bos tauru
42	16.8	84.0	10	AL669954	AL669954 Mouse DNA
43	16.8	84.0	2	AC141138	AC141138 Rattus no
44	16.8	84.0	8	AC148290	AC148290 Medicago
45	16.8	84.0	2	AC151300	AC151300 Xenopus t
46	16.8	84.0	2	AC145410	AC145410 Bos tauru
47	16.8	84.0	9	AB020868	AB020868 Homo sapi
48	16.8	84.0	9	AC116903	AC116903 Homo sapi
49	16.8	84.0	9	AC116162	AC116162 Homo sapi
50	16.8	84.0	9	AP006203	AP006203 Homo sapi
51	16.8	84.0	9	AC099019	AC099019 Drosophill
52	16.8	84.0	2	AC141032	AC141032 Rattus no
53	16.8	84.0	9	AP006208	AP006208 Homo sapi
54	16.8	84.0	2	AC022372	AC022372 Homo sapi
55	16.8	84.0	3	AC008348	AC008348 Drosophill
56	16.8	84.0	2	AC149202	AC149202 Macaca mu
57	16.8	84.0	2	AC137548	AC137548 Macaca mu
58	16.8	84.0	10	AC125151	AC125151 Mus muscu
59	16.8	84.0	2	AC135719	AC135719 Homo sapi
60	16.8	84.0	9	AC124242	AC124242 Homo sapi
61	16.8	84.0	2	AC126953	AC126953 Rattus no
62	16.8	84.0	2	AC109879	AC109879 Rattus no
63	16.8	84.0	2	AC096624	AC096624 Mus muscu
64	16.8	84.0	2	AC099108	AC099108 Rattus no
65	16.8	84.0	2	AC121224	AC121224 Rattus no
66	16.8	84.0	2	AC118336	AC118336 Rattus no
67	16.8	84.0	2	AC126485	AC126485 Rattus no
68	16.8	84.0	2	AC123483	AC123483 Rattus no
69	16.8	84.0	2	AC111820	AC111820 Rattus no
70	16.8	84.0	2	AC126196	AC126196 Rattus no
71	16.8	84.0	2	AC095304	AC095304 Rattus no
72	16.8	84.0	2	AC122970	AC122970 Rattus no
73	16.8	84.0	2	AE003456	AE003456 Drosophill
74	16.8	84.0	11	BV073605	BV073605 S212P6673
75	16.4	82.0	3	DPADHGN57	X62237 D.pseudosona
76	16.4	82.0	3	TCU22317	TCU22317 Trypanosona
77	16.4	82.0	9	AL139827	AL139827 Human DNA
78	16.4	82.0	2	AC101057	AC101057 Mus muscu
79	16.4	82.0	2	AC084171	AC084171 Homo sapi
80	16.4	82.0	2	AC013789	AC013789 Homo sapi
81	16.4	82.0	10	AC104827	AC104827 Homo sapi
82	16.4	82.0	9	AC012511	AC012511 Homo sapi
83	16.4	82.0	2	AC129010	AC129010 Leishmani
84	16.4	82.0	2	CR735113	CR735113 Danio rer
85	16.4	82.0	2	AC131800	AC131800 Mus muscu
86	16.4	82.0	1	AC012288	AC012288 Homo sapi
87	16.4	82.0	2	AC021566	AC021566 Homo sapi
88	16.4	82.0	2	AC116622	AC116622 Rattus no
89	16.4	82.0	2	AC102757	AC102757 Mus muscu
90	16.4	82.0	10	AC099520	AC099520 Homo sapi
91	16.4	82.0	9	AC099520	AC099520 Homo sapi
92	16.4	82.0	9	AC099520	AC099520 Homo sapi

93	16.4	82.0	181585	9	AL139350	AL139350 Human DNA	166	15.8	79.0	1558	3	AY089564	AY089564 Drosophila
94	16.4	82.0	189267	2	AC137230	Rattus no	167	15.8	79.0	2269	9	BC017299	BC017299 Homo sapi
95	16.4	82.0	196602	2	AC114248	Rattus no	168	15.8	79.0	2269	9	AY026763	AY026763 Homo sapi
96	16.4	82.0	208595	2	AC139990	Rattus no	c 169	15.8	79.0	2284	5	BC072060	BC072060 Xenopus l
97	16.4	82.0	206741	10	AL7332547	AL7332547 Mouse DNA	c 170	15.8	79.0	2496	5	XLPOU3	X64835 X.laevis xl
98	16.4	82.0	213684	10	AC132616	AC132616 Mus muscu	c 171	15.8	79.0	2526	5	BC041298	BC041298 Xenopus l
99	16.4	82.0	213903	2	AC111731	AC111731 Rattus no	c 172	15.8	79.0	2927	5	BC080492	BC080492 Xenopus t
100	16.4	82.0	215530	2	AC105874	AC105874 Rattus no	c 173	15.8	79.0	3181	10	MMJ4740	AJ000740 Mus Muscu
101	16.4	82.0	223384	2	AC137440	AC137440 Rattus no	c 174	15.8	79.0	3186	9	AK123944	AK123944 Homo sapi
102	16.4	82.0	223693	2	AC103053	AC103053 Rattus no	175	15.8	79.0	3364	10	BC007130	BC007130 Mus muscu
103	16.4	82.0	225014	2	AC096045	AC096045 Rattus no	176	15.8	79.0	3545	6	CQ583988	CQ583988 Sequence
104	16.4	82.0	226604	2	AC094308	AC094308 Rattus no	c 177	15.8	79.0	4122	5	XLPOU3B	X96423 X.laevis PO
105	16.4	82.0	232843	2	AC0943578	AC0943578 Mus muscu	c 178	15.8	79.0	4282	10	MMAC13A	M06640 Mouse mRNA
106	16.4	82.0	243502	10	AL772402	AL772402 Mouse DNA	179	15.8	79.0	4409	3	AY157180	AY157180 Caryophyl
107	16.4	82.0	253401	2	AC097127	AC097127 Rattus no	c 180	15.8	79.0	5303	9	AC079143	AC079143 Homo sapi
108	16.4	82.0	266634	2	AC123487	AC123487 Rattus no	181	15.8	79.0	5856	6	CQ583994	CQ583994 Sequence
109	16.4	82.0	310003	9	AF178030	AF178030 Homo sapi	182	15.8	79.0	6413	3	AF047351	AF047351 Leishmani
110	16.4	82.0	347286	2	AC102740	AC102740 Mus muscu	183	15.8	79.0	15921	3	AF034856	AF034856 Drosophila
111	16	80.0	368	6	AX8933238	AX8933238 Sequence	184	15.8	79.0	24337	2	AC013130	AC013130 Drosophila
112	16	80.0	368	6	BD028771	BD028771 Sequence	185	15.8	79.0	26467	9	AP001472	AP001472 Homo sapi
113	16	80.0	50865	9	AL929444	AL929444 Human DNA	c 186	15.8	79.0	33931	9	AC022156	AC022156 Homo sapi
114	16	80.0	69206	2	AC020210	AC020210 Drosophila	187	15.8	79.0	36885	2	AC011015	AC011015 Leishmani
115	16	80.0	80449	2	AC114714	AC114714 Rattus no	188	15.8	79.0	36881	2	AC0114041	AC0114041 Drosophila
116	16	80.0	84350	2	AC016861	AC016861 Homo sapi	189	15.8	79.0	38817	10	AF427516	AF427516 Mus muscu
117	16	80.0	113810	9	AC068292	AC068292 Homo sapi	c 190	15.8	79.0	39855	9	AC020954	AC020954 Homo sapi
118	16	80.0	164689	2	AL953873	AL953873 Homo sapi	c 191	15.8	79.0	46064	5	AF461063	AF461063 Takifugu
119	16	80.0	168315	10	AC124451	AC124451 Mus muscu	192	15.8	79.0	48554	9	AC133104	AC133104 Homo sapi
120	16	80.0	168460	2	AC011585	AC011585 Homo sapi	193	15.8	79.0	58569	2	AC015033	AC015033 Drosophila
121	16	80.0	174274	2	AX957252	AX957252 Danio rer	194	15.8	79.0	59351	2	AC100656	AC100656 Mus muscu
122	16	80.0	174274	2	AX957252	AX957252 Danio rer	195	15.8	79.0	61170	2	AC068163	AC068163 Homo sapi
123	16	80.0	175436	2	AC122986	AC122986 Rattus no	c 196	15.8	79.0	61170	2	AC068163	AC068163 Homo sapi
124	16	80.0	175438	10	AC116598	AC116598 Mus muscu	197	15.8	79.0	67307	2	AC120598	Continuation (4 of
125	16	80.0	176195	3	AC012165	AC012165 Drosophila	c 198	15.8	79.0	69861	2	AC032033	Continuation (4 of
126	16	80.0	177911	2	AC021362	AC021362 Homo sapi	c 199	15.8	79.0	78456	10	AL627393	AL627393 Homo sapi
127	16	80.0	178810	9	AC010653	AC010653 Homo sapi	200	15.8	79.0	82827	9	AC079395	AC079395 Homo sapi
128	16	80.0	180047	2	AC105174	AC105174 Mus muscu	201	15.8	79.0	82993	6	AX704274	AX704274 Sequence
129	16	80.0	182912	9	AC068060	AC068060 Homo sapi	202	15.8	79.0	91231	9	AL157364	AL157364 Human DNA
130	16	80.0	182913	2	AC068556	AC068556 Homo sapi	203	15.8	79.0	91489	9	AC010444	AC010444 Homo sapi
131	16	80.0	201862	2	AC150069	AC150069 Gallus ga	204	15.8	79.0	91825	2	AC019854	AC019854 Drosophila
132	16	80.0	203086	10	AC099948	AC099948 Mus muscu	c 205	15.8	79.0	97017	10	AL845423	AL845423 Mouse DNA
133	16	80.0	207432	3	AE003513	AE003513 Drosophila	c 206	15.8	79.0	98056	9	AC006061	AC006061 Homo sapi
134	16	80.0	220458	9	AC010320	AC010320 Homo sapi	207	15.8	79.0	110000	1	AE000516	Continuation (22 o
135	16	80.0	223708	2	AC102724	AC102724 Mus muscu	c 208	15.8	79.0	110000	2	AC008576	Continuation (2 of
136	16	80.0	239456	2	AC108995	AC108995 Rattus no	c 209	15.8	79.0	110000	2	AC079517	Continuation (3 of
137	16	80.0	243310	2	AC096208	AC096208 Rattus no	c 210	15.8	79.0	110000	2	AC132794	Continuation (4 of
138	16	80.0	245082	2	AC097237	AC097237 Rattus no	211	15.8	79.0	110000	3	AE003426	AE003426 Drosophila
139	16	80.0	249310	2	AC096891	AC096891 Rattus no	c 212	15.8	79.0	110000	8	CR382129	Continuation (9 of
140	16	80.0	257862	2	AC098100	AC098100 Rattus no	c 213	15.8	79.0	111061	9	AC078942	AC078942 Homo sapi
141	16	80.0	258988	2	AC103322	AC103322 Rattus no	214	15.8	79.0	114000	2	AC125412	AC125412 Leishmani
142	16	80.0	270456	2	AC111575	AC111575 Rattus no	c 215	15.8	79.0	116908	2	AC018365	AC018365 Mus muscu
143	15.8	79.0	328	11	BV074663	BV074663 S208P6019	216	15.8	79.0	123244	10	BX004791	BX004791 Mouse DNA
144	15.8	79.0	336	11	MMST5162	236700 M.musculus	217	15.8	79.0	127883	10	AC114434	AC114434 Rattus no
145	15.8	79.0	430	4	BTMSAT76	D32088 Octopus vul	c 218	15.8	79.0	128614	10	BX510300	BX510300 Mouse DNA
146	15.8	79.0	442	3	OTCSINEF	D32088 Octopus vul	c 219	15.8	79.0	131000	9	AP001017	AP001017 Homo sapi
147	15.8	79.0	458	4	BTMSAT72	D32094 Octopus vul	220	15.8	79.0	131264	2	AC104198	AC104198 Mus muscu
148	15.8	79.0	460	3	OTCSINEL	D32094 Octopus vul	221	15.8	79.0	131819	9	HS529N6	AL034376 Human DNA
149	15.8	79.0	471	9	HSTCB4	X74844 H.sapiens (	c 222	15.8	79.0	132176	2	AC150122	AC150122 Gallus ga
150	15.8	79.0	609	9	HS3339102	AJ339102 Homo sapi	c 223	15.8	79.0	135276	2	AC073603	AC073603 Mus muscu
151	15.8	79.0	628	11	BV021838	BV021838 S212P6375	224	15.8	79.0	136010	2	AC150128	AC150128 Gallus ga
152	15.8	79.0	642	9	HS3339311	AJ339311 Homo sapi	c 225	15.8	79.0	137286	4	AL773527	AL773527 Pig DNA s
153	15.8	79.0	642	9	HS3339313	AJ339313 Homo sapi	c 226	15.8	79.0	137898	2	AC140552	AC140552 Mus muscu
154	15.8	79.0	656	9	HS3342373	AJ342373 Homo sapi	c 227	15.8	79.0	138048	2	AC113575	AC113575 Tetraodon
155	15.8	79.0	662	9	HS332969	AJ322969 Homo sapi	c 228	15.8	79.0	138217	10	AL669905	AL669905 Mouse DNA
156	15.8	79.0	665	9	HS3342510	AJ342510 Homo sapi	c 229	15.8	79.0	138320	2	AC150081	AC150081 Gallus ga
157	15.8	79.0	667	9	HS3329574	AJ329574 Homo sapi	c 230	15.8	79.0	139894	2	AC151208	AC151208 Bos tauru
158	15.8	79.0	678	9	HS3340234	AJ340234 Homo sapi	c 231	15.8	79.0	141073	9	AC105108	AC105108 Homo sapi
159	15.8	79.0	682	9	HS3340223	AJ340223 Homo sapi	c 232	15.8	79.0	141766	2	AC148652	AC148652 Medicago
160	15.8	79.0	729	9	HS3330600	AJ330600 Homo sapi	233	15.8	79.0	143895	2	CR407557	CR407557 Danio rer
161	15.8	79.0	738	9	HS3338205	AJ338205 Homo sapi	c 234	15.8	79.0	144699	2	AC125881	AC125881 Rattus no
162	15.8	79.0	765	5	CR391089	CR391089 Gallus ga	c 235	15.8	79.0	145709	2	AC069288	AC069288 Homo sapi
163	15.8	79.0	931	3	OTCSINEA	D32083 Octopus vul	c 236	15.8	79.0	147634	9	AC055740	AC055740 Homo sapi
164	15.8	79.0	1105	9	BC063580	BC063580 Homo sapi	c 237	15.8	79.0	149427	2	AC121385	AC121385 Rattus no
165	15.8	79.0	1217	9	AY177407	AY177407 Homo sapi	c 238	15.8	79.0	150069	2	AC041042	AC041042 Homo sapi

C 239	15.8	79.0	151302	9	AC073617	AC073617 Homo sapi	312	15.8	79.0	182553	2	AC149646	AC149646 Bos tauru
C 240	15.8	79.0	152827	2	AC143347	AC143347 Homo sapi	313	15.8	79.0	182661	10	AC121602	AC121602 Mus muscu
C 241	15.8	79.0	153565	10	AC080016	AC080016 Mus muscu	C 314	15.8	79.0	182669	10	AC135859	AC135859 Mus muscu
C 242	15.8	79.0	154154	2	AC121679	AC121679 Rattus no	C 315	15.8	79.0	183239	9	AC015906	AC015906 Homo sapi
C 243	15.8	79.0	155001	2	AL356474	AL356474 Homo sapi	C 316	15.8	79.0	183273	9	AC093898	AC093898 Homo sapi
C 244	15.8	79.0	155670	3	AC020297	AC020297 Drosophill	C 317	15.8	79.0	183586	9	AC092562	AC092562 Papio ham
C 245	15.8	79.0	155685	3	AC098575	AC098575 Drosophill	C 318	15.8	79.0	183720	2	AC092912	AC092912 Homo sapi
C 246	15.8	79.0	156684	10	AC058787	AC058787 Mus muscu	C 319	15.8	79.0	183897	2	AC150665	AC150665 Bos tauru
C 247	15.8	79.0	157091	2	AC027772	AC027772 Homo sapi	C 320	15.8	79.0	184163	2	AC149723	AC149723 Bos tauru
C 248	15.8	79.0	157435	5	AL935029	AL935029 Zebrafish	C 321	15.8	79.0	184798	2	AC139981	AC139981 Rattus no
C 249	15.8	79.0	157633	9	AC008687	AC008687 Homo sapi	C 322	15.8	79.0	185800	2	AC135224	AC135224 Gallus ga
C 250	15.8	79.0	157669	2	AC143345	AC143345 Homo sapi	C 323	15.8	79.0	186058	2	AC112245	AC112245 Homo sapi
C 251	15.8	79.0	157733	2	AC055810	AC055810 Homo sapi	C 324	15.8	79.0	186144	2	AC069048	AC069048 Homo sapi
C 252	15.8	79.0	157812	9	AC078732	AC078732 Homo sapi	C 325	15.8	79.0	186258	5	BX640480	BX640480 Zebrafish
C 253	15.8	79.0	157921	2	AC027726	AC027726 Homo sapi	C 326	15.8	79.0	186347	2	AC107710	AC107710 Mus muscu
C 254	15.8	79.0	158121	2	AC147848	AC147848 Ocolemur	C 327	15.8	79.0	186748	3	AC012163	AC012163 Drosophill
C 255	15.8	79.0	159377	2	AC147522	AC147522 Ocolemur	C 328	15.8	79.0	187058	10	AC126410	AC126410 Mus muscu
C 256	15.8	79.0	161354	2	AC113374	AC113374 Homo sapi	C 329	15.8	79.0	187343	2	AC016888	AC016888 Homo sapi
C 257	15.8	79.0	162835	10	AC121866	AC121866 Mus muscu	C 330	15.8	79.0	187344	9	AL354873	AL354873 Human DNA
C 258	15.8	79.0	163724	2	CR788299	CR788299 Danio rer	C 331	15.8	79.0	187450	10	BX322546	BX322546 Mouse DNA
C 259	15.8	79.0	163923	2	AC074278	AC074278 Homo sapi	C 332	15.8	79.0	188646	2	AC141652	AC141652 Rattus no
C 260	15.8	79.0	164511	9	AC004894	AC004894 Homo sapi	C 333	15.8	79.0	189023	9	AL592114	AL592114 Human DNA
C 261	15.8	79.0	164635	2	AC143346	AC143346 Homo sapi	C 334	15.8	79.0	189263	10	AC134595	AC134595 Mus muscu
C 262	15.8	79.0	165002	2	AC150642	AC150642 Bos tauru	C 335	15.8	79.0	189650	10	AC136091	AC136091 Rattus no
C 263	15.8	79.0	165033	2	AC087125	AC087125 Mus muscu	C 336	15.8	79.0	191674	4	CR536601	CR536601 Platypus
C 264	15.8	79.0	165212	10	AL596125	AL596125 Mouse DNA	C 337	15.8	79.0	191923	9	AC109480	AC109480 Homo sapi
C 265	15.8	79.0	165378	2	AC143332	AC143332 Homo sapi	C 338	15.8	79.0	192015	2	AC119223	AC119223 Mus muscu
C 266	15.8	79.0	166065	2	AC149014	AC149014 Papio anu	C 339	15.8	79.0	192556	2	AC103942	AC103942 Mus muscu
C 267	15.8	79.0	166171	2	AL359550	AL359550 Homo sapi	C 340	15.8	79.0	192569	2	AC133928	AC133928 Lemur cat
C 268	15.8	79.0	166787	9	AC055723	AC055723 Homo sapi	C 341	15.8	79.0	192687	2	AC129244	AC129244 Rattus no
C 269	15.8	79.0	167053	10	AC124447	AC124447 Mus muscu	C 342	15.8	79.0	192994	2	CR318588	CR318588 Danio rer
C 270	15.8	79.0	167905	9	AC116606	AC116606 Homo sapi	C 343	15.8	79.0	193147	2	AC149200	AC149200 Macaca mu
C 271	15.8	79.0	168277	2	AC147934	AC147934 Ocolemur	C 344	15.8	79.0	193286	2	AC110119	AC110119 Rattus no
C 272	15.8	79.0	168509	2	AC115344	AC115344 Rattus no	C 345	15.8	79.0	193580	2	AC117284	AC117284 Rattus no
C 273	15.8	79.0	168671	9	AC019247	AC019247 Homo sapi	C 346	15.8	79.0	194109	10	AC121819	AC121819 Mus muscu
C 274	15.8	79.0	168779	9	AC026405	AC026405 Homo sapi	C 347	15.8	79.0	194171	2	AC126452	AC126452 Mus muscu
C 275	15.8	79.0	169046	5	AL928828	AL928828 Zebrafish	C 348	15.8	79.0	195728	2	CR384051	CR384051 Danio rer
C 276	15.8	79.0	169155	5	BX088685	BX088685 Zebrafish	C 349	15.8	79.0	197229	2	AC021162	AC021162 Homo sapi
C 277	15.8	79.0	169457	5	AC009216	AC009216 Drosophill	C 350	15.8	79.0	197679	10	AC128663	AC128663 Mus muscu
C 278	15.8	79.0	170102	9	AC009237	AC009237 Homo sapi	C 351	15.8	79.0	198116	2	AC123155	AC123155 Rattus no
C 279	15.8	79.0	170323	10	AC124459	AC124459 Mus muscu	C 352	15.8	79.0	198919	10	AC140277	AC140277 Mus muscu
C 280	15.8	79.0	171451	10	AC079444	AC079444 Mus muscu	C 353	15.8	79.0	198996	2	AC134845	AC134845 Mus muscu
C 281	15.8	79.0	171557	2	AC147276	AC147276 Pan trogl	C 354	15.8	79.0	200946	2	AC148421	AC148421 Meleagris
C 282	15.8	79.0	171961	9	AC113408	AC113408 Homo sapi	C 355	15.8	79.0	200985	3	AC010060	AC010060 Drosophill
C 283	15.8	79.0	171982	2	AC120916	AC120916 Rattus no	C 356	15.8	79.0	201485	2	AC134402	AC134402 Mus muscu
C 284	15.8	79.0	172253	2	CR847922	CR847922 Danio rer	C 357	15.8	79.0	201602	2	AC074146	AC074146 Mus muscu
C 285	15.8	79.0	172398	2	AC055121	AC055121 Homo sapi	C 358	15.8	79.0	202083	2	AC023833	AC023833 Mus muscu
C 286	15.8	79.0	172382	2	AC149046	AC149046 Papio anu	C 359	15.8	79.0	202086	10	AC080015	AC080015 Mus muscu
C 287	15.8	79.0	172963	9	AC100763	AC100763 Homo sapi	C 360	15.8	79.0	202151	10	AC102358	AC102358 Mus muscu
C 288	15.8	79.0	173126	9	AC005988	AC005988 Homo sapi	C 361	15.8	79.0	203495	2	BX936436	BX936436 Danio rer
C 289	15.8	79.0	173667	2	AC134316	AC134316 Lemur cat	C 362	15.8	79.0	203765	9	AC087289	AC087289 Homo sapi
C 290	15.8	79.0	173882	5	BX640462	BX640462 Zebrafish	C 363	15.8	79.0	203829	5	AC144485	AC144485 Gasterost
C 291	15.8	79.0	175231	10	AC034108	AC034108 Mus muscu	C 364	15.8	79.0	204214	10	AL672251	AL672251 Mouse DNA
C 292	15.8	79.0	176384	2	AC150507	AC150507 Bos tauru	C 365	15.8	79.0	204239	10	AC114819	AC114819 Mus muscu
C 293	15.8	79.0	176767	9	AF196779	AF196779 Homo sapi	C 366	15.8	79.0	204892	2	AC151244	AC151244 Bos tauru
C 294	15.8	79.0	177956	2	AC110916	AC110916 Mus muscu	C 367	15.8	79.0	205503	2	AC148001	AC148001 Mus muscu
C 295	15.8	79.0	178208	2	AC136432	AC136432 Homo sapi	C 368	15.8	79.0	205553	2	AC118516	AC118516 Rattus no
C 296	15.8	79.0	178346	2	AC134539	AC134539 Mus muscu	C 369	15.8	79.0	205985	2	AC145399	AC145399 Rattus no
C 297	15.8	79.0	178443	10	AL592185	AL592185 Mouse DNA	C 370	15.8	79.0	207424	10	AC090489	AC090489 Genomic s
C 298	15.8	79.0	178457	5	AL935031	AL935031 Zebrafish	C 371	15.8	79.0	208214	10	AL645908	AL645908 Mouse DNA
C 299	15.8	79.0	178728	2	AC143348	AC143348 Homo sapi	C 372	15.8	79.0	209877	9	AC012509	AC012509 Homo sapi
C 300	15.8	79.0	179145	9	AC009785	AC009785 Homo sapi	C 373	15.8	79.0	210206	2	AC139598	AC139598 Rattus no
C 301	15.8	79.0	179155	9	AL353748	AL353748 Human DNA	C 374	15.8	79.0	210206	2	AC139598	AC139598 Rattus no
C 302	15.8	79.0	179192	2	AC135529	AC135529 Rattus no	C 375	15.8	79.0	211350	2	AC112604	AC112604 Rattus no
C 303	15.8	79.0	179501	9	AL591543	AL591543 Human DNA	C 376	15.8	79.0	211763	10	AC123558	AC123558 Mus muscu
C 304	15.8	79.0	179592	2	AC009723	AC009723 Homo sapi	C 377	15.8	79.0	212120	2	AC023414	AC023414 Homo sapi
C 305	15.8	79.0	179649	3	AC010045	AC010045 Drosophill	C 378	15.8	79.0	212628	2	AC097283	AC097283 Rattus no
C 306	15.8	79.0	179711	2	AC019312	AC019312 Homo sapi	C 379	15.8	79.0	213746	2	CR847568	CR847568 Danio rer
C 307	15.8	79.0	179941	2	AC150586	AC150586 Bos tauru	C 380	15.8	79.0	214527	9	CNS00007	AL049819 Human chr
C 308	15.8	79.0	180787	3	AC009375	AC009375 Drosophill	C 381	15.8	79.0	215196	2	AC073796	AC073796 Mus muscu
C 309	15.8	79.0	180914	9	CNS01051	AL121612 Human chr	C 382	15.8	79.0	215199	2	AC149052	AC149052 Mus muscu
C 310	15.8	79.0	181111	10	AC129301	AC129301 Mus muscu	C 383	15.8	79.0	215863	2	AC136812	AC136812 Rattus no
C 311	15.8	79.0	181162	5	AC144487	AC144487 Gasterost	C 384	15.8	79.0	215914	2	AC127875	AC127875 Rattus no

c 385	15.8	79.0	216807	10	AC079681	Mus muscu
c 386	15.8	79.0	219121	2	AC134718	Rattus no
c 387	15.8	79.0	219409	2	AC134858	Mus muscu
c 388	15.8	79.0	219747	2	AC098304	Rattus no
c 389	15.8	79.0	220788	2	AC130904	Rattus no
c 390	15.8	79.0	220805	2	AC118515	Rattus no
c 391	15.8	79.0	220931	2	AC116066	Rattus no
c 392	15.8	79.0	220931	2	AC131405	Rattus no
c 393	15.8	79.0	223344	2	AC106193	Rattus no
c 394	15.8	79.0	223345	2	AC123324	Rattus no
c 395	15.8	79.0	223971	2	AC115753	Mus muscu
c 396	15.8	79.0	223979	2	AC096151	Rattus no
c 397	15.8	79.0	224592	2	AC074209	Mus muscu
c 398	15.8	79.0	224946	2	AC115413	Rattus no
c 399	15.8	79.0	225005	2	AC107607	Rattus no
c 400	15.8	79.0	225274	2	AC130251	Rattus no
c 401	15.8	79.0	225924	2	AC074210	Mus muscu
c 402	15.8	79.0	226524	2	AC097833	Rattus no
c 403	15.8	79.0	226665	2	AC098153	Rattus no
c 404	15.8	79.0	227364	2	AC113732	Rattus no
c 405	15.8	79.0	227833	2	AC128446	Rattus no
c 406	15.8	79.0	227834	4	AC150561	Bos tauru
c 407	15.8	79.0	228858	2	AC120623	Rattus no
c 408	15.8	79.0	229063	2	AC147751	Otolemur
c 409	15.8	79.0	229131	2	AC114701	Rattus no
c 410	15.8	79.0	229482	2	AC074149	Mus muscu
c 411	15.8	79.0	230783	2	AC118913	Rattus no
c 412	15.8	79.0	231234	2	AC148222	Colobus g
c 413	15.8	79.0	231722	2	AC097596	Rattus no
c 414	15.8	79.0	231864	2	AC129642	Rattus no
c 415	15.8	79.0	232473	10	AC128896	Mus muscu
c 416	15.8	79.0	233566	2	AC118900	Rattus no
c 417	15.8	79.0	233784	2	AC130877	Rattus no
c 418	15.8	79.0	234399	2	AC117147	Rattus no
c 419	15.8	79.0	234655	2	AC096021	Rattus no
c 420	15.8	79.0	234681	2	AC095940	Rattus no
c 421	15.8	79.0	234842	2	AC127153	Rattus no
c 422	15.8	79.0	234967	2	AC134282	Rattus no
c 423	15.8	79.0	235025	2	AC114149	Rattus no
c 424	15.8	79.0	235261	2	AC115280	Rattus no
c 425	15.8	79.0	236456	2	AC036147	Mus muscu
c 426	15.8	79.0	237327	2	AC103454	Rattus no
c 427	15.8	79.0	238050	2	AC094265	Rattus no
c 428	15.8	79.0	239570	2	AC137983	Mus muscu
c 429	15.8	79.0	239946	2	AC103505	Rattus no
c 430	15.8	79.0	240142	2	AC134295	Rattus no
c 431	15.8	79.0	240596	2	AC095941	Rattus no
c 432	15.8	79.0	241069	2	AC124894	Rattus no
c 433	15.8	79.0	241721	2	AC111729	Rattus no
c 434	15.8	79.0	242465	2	AC128332	Rattus no
c 435	15.8	79.0	243354	5	BX470172	Zebrafish
c 436	15.8	79.0	243537	2	AC132629	Rattus no
c 437	15.8	79.0	243981	2	AC111281	Rattus no
c 438	15.8	79.0	244068	2	AC131345	Rattus no
c 439	15.8	79.0	244377	2	AC110661	Rattus no
c 440	15.8	79.0	245153	2	AC119557	Rattus no
c 441	15.8	79.0	245417	2	AC105891	Rattus no
c 442	15.8	79.0	245737	2	AC122654	Rattus no
c 443	15.8	79.0	246006	2	AC102141	Mus muscu
c 444	15.8	79.0	246772	2	AC112469	Rattus no
c 445	15.8	79.0	247046	2	AC119441	Rattus no
c 446	15.8	79.0	247146	2	AC134757	Rattus no
c 447	15.8	79.0	248196	2	AC127847	Rattus no
c 448	15.8	79.0	248294	2	AC097690	Rattus no
c 449	15.8	79.0	249364	2	AC098332	Rattus no
c 450	15.8	79.0	249549	2	AC109236	Mus muscu
c 451	15.8	79.0	249781	2	AC098954	Rattus no
c 452	15.8	79.0	250529	2	AC097607	Rattus no
c 453	15.8	79.0	251478	2	AC105849	Rattus no
c 454	15.8	79.0	251495	2	AC113722	Rattus no
c 455	15.8	79.0	251740	2	AC123460	Rattus no
c 456	15.8	79.0	252041	2	AC114432	Rattus no
c 457	15.8	79.0	252257	2	AC096047	Rattus no

c 458	15.8	79.0	252580	2	AC121227	Rattus no
c 459	15.8	79.0	252775	2	AC109711	Rattus no
c 460	15.8	79.0	253922	2	AC111987	Rattus no
c 461	15.8	79.0	254380	2	AC105841	Rattus no
c 462	15.8	79.0	254380	10	AC118234	Mus muscu
c 463	15.8	79.0	256220	2	AC109981	Rattus no
c 464	15.8	79.0	257333	2	AC132722	Rattus no
c 465	15.8	79.0	257396	2	AC130569	Rattus no
c 466	15.8	79.0	257645	2	AC103094	Rattus no
c 467	15.8	79.0	257645	2	AC116230	Rattus no
c 468	15.8	79.0	262581	2	AC095238	Rattus no
c 469	15.8	79.0	263939	2	AC097113	Rattus no
c 470	15.8	79.0	265067	2	AC110852	Rattus no
c 471	15.8	79.0	265515	2	AC109853	Rattus no
c 472	15.8	79.0	266762	2	AC109853	Rattus no
c 473	15.8	79.0	271034	2	AC102996	Rattus no
c 474	15.8	79.0	271496	2	AC142073	Rattus no
c 475	15.8	79.0	272535	5	EX005212	Zebrafish
c 476	15.8	79.0	275277	2	AC129438	Rattus no
c 477	15.8	79.0	277347	2	AC130968	Rattus no
c 478	15.8	79.0	281117	2	AC110645	Rattus no
c 479	15.8	79.0	289140	3	AE003533	Drosophil
c 480	15.8	79.0	290138	2	AC097202	Rattus no
c 481	15.8	79.0	291050	1	EX248340	Mycobacte
c 482	15.8	79.0	295466	2	AC123317	Rattus no
c 483	15.8	79.0	299804	3	AE003507	Drosophil
c 484	15.8	79.0	302050	3	AE003522	Drosophil
c 485	15.8	79.0	313499	2	AC119366	Rattus no
c 486	15.8	79.0	318535	2	AC107011	Rattus no
c 487	15.8	79.0	327876	2	AC095476	Rattus no
c 488	15.8	79.0	327876	2	AC095476	Rattus no
c 489	15.8	79.0	328863	2	AL627202	Homo sapi
c 490	15.8	79.0	336392	2	AC130441	Rattus no
c 491	15.8	79.0	340000	9	AP001759	Homo sapi
c 492	15.8	79.0	342336	1	EX842578	Mycobacte
c 493	15.4	77.0	286	11	AU049634	AU049694 Rattus no
c 494	15.4	77.0	334	11	G47212	G47212 Z15893_1 Ze
c 495	15.4	77.0	351	5	AY096855	AY096855 Glyphorin
c 496	15.4	77.0	351	5	AY096871	AY096871 Glyphorin
c 497	15.4	77.0	351	5	AY096872	AY096872 Glyphorin
c 498	15.4	77.0	351	5	AY096873	AY096873 Glyphorin
c 499	15.4	77.0	351	5	AY096874	AY096874 Glyphorin
c 500	15.4	77.0	351	5	AY096875	AY096875 Glyphorin

ALIGNMENTS

RESULT 1	ARI146422/c	Sequence 39 from patent US 6218510.	124 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	ARI146422					
DEFINITION	ARI146422.1	GI:15109611				
ACCESSION	ARI146422.1					
VERSION	ARI146422.1					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 124)					
AUTHORS	Sharpe,A.H., Borriello,F., Freeman,G.J. and Nadler,L.M.					
TITLE	B7-1 and B7-2 polypeptides					
JOURNAL	Patent: US 6218510-A 39 17-APR-2001;					
FEATURES	Location/Qualifiers					
source	1..124					
	/organism="unknown"					
	/mol_type="unassigned DNA"					
ORIGIN						
Query Match	100.0%;	Score 20;	DB 6;	Length 124;		
Best Local Similarity	100.0%;	Pred. No. 44;				
Matches	20;	Conservative	0;	Mismatches	0;	Gaps 0;
QY	1	CCTGTGTCGTGCTAGTCCC	20			



```

Db      57  CGTGTCTGTGCTAGTCC 38
|||||
RESULT 2
AR381504/c
LOCUS      AR381504      124 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 39 from patent US 6608180.
ACCESSION  AR381504
VERSION     AR381504.1  GI:40089565
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 124)
AUTHORS   Sharpe,A.H., Borriello,F., Freeman,G.J. and Nadler,L.M.
TITLE     B7-specific antibodies
JOURNAL   Patent: US 6608180-A 39 19-AUG-2003;
FEATURES  Location/Qualifiers
source    1..124
           /organism="unknown"
           /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  1  CGTGTCTGTGCTAGTCC 20
      |||||||
Db   57  CGTGTCTGTGCTAGTCC 38
|||||
RESULT 3
AX884562/c
LOCUS      AX884562      430 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 425 from Patent EP1033401.
ACCESSION  AX884562
VERSION     AX884562.1  GI:40039847
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE     Expressed sequence tags and encoded human proteins
JOURNAL   Patent: EP 1033401-A 425 06-SEP-2000;
          Genset (FR)
FEATURES  Location/Qualifiers
source    1..430
           /organism="Homo sapiens"
           /mol_type="unassigned DNA"
           /db_xref="taxon:9606"
           125..430
           /note="unnamed protein product"
           /codon_start=1
           /protein_id="CAE98676.1"
           /db_xref="GI:40039848"
           /translation="MDPQCTMGLNLLFWAFLLSGAAPLKIOAFNFTADLPQCFAN
           "
           SQNQSLSELVVFVQDOENLVLEVLGKFKDSVHSKYMGRTSPDSDSWTLRLHLQI
sig_peptide 125..199
           /note="score 7.1 seq LFVMAFLLSGAAP/LK"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  1  CGTGTCTGTGCTAGTCC 20
      |||||||
Db   75  CGTGTCTGTGCTAGTCC 56
|||||
RESULT 4
BD024172/c
LOCUS      BD024172      430 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION  BD024172
VERSION     BD024172.1  GI:22565395
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 430)
AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE     Sequence tag and encoded human protein
JOURNAL   Patent: JP 2001269182-A 418 02-OCT-2001;
          GENSET
COMMENT    OS Homo sapiens (human)
          PN JP 2001269182-A/418
          PD 02-OCT-2001
          PF 24-FEB-2000 JP 2000118773
          PR 26-FEB-1999 US 60/122487
          PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
          PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
          C12N5/10,
          PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
          G06F15/40
          CC score 7.1
          CC seq LFVMAFLLSGAAP/LK
          FH Key Location/Qualifiers
          FT CDS 125..430
          FT sig_peptide 125..199.
          Location/Qualifiers
          1..430
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  1  CGTGTCTGTGCTAGTCC 20
      |||||||
Db   75  CGTGTCTGTGCTAGTCC 56
|||||
RESULT 5
HSB72S1/c
LOCUS      HSB72S1      478 bp      DNA      linear      PRI 02-FEB-1996
DEFINITION Human CTLA-4 counter-receptor B7.2 (B7.2) gene, exon 1.
ACCESSION  U17715
VERSION     U17715.1  GI:808025
KEYWORDS   1 of 8
SEGMENT    Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 478)
          Jellis,C.L., Wang,S.S., Rennert,P., Borriello,F., Sharpe,A.H.,
          Green,N.R. and Gray,G.S.
          Genomic organization of the gene coding for the costimulatory human
          B-lymphocyte antigen B7-2 (CD86)
          Immunogenetics 42 (2), 85-89 (1995)
          9531831
          MEDLINE
          PUBMED 7541777
          2 (bases 173 to 292)
          Freeman,G.J., Gribben,J.G., Boussiotis,V.A., Ng,J.W., Restivo,V.A.
          Jr., Lombard,L.A., Gray,G.S. and Nadler,L.M.

```

TITLE Cloning of B7-2: a CTLA-4 counter-receptor that costimulates human T cell proliferation

JOURNAL Science 262 (5135), 909-911 (1993)

MEDLINE 94053735

PUBMED 7694363

REFERENCE 3 (sites)

AUTHORS Azuma,M., Ito,D., Yagita,H., Okumura,K., Phillips,J.H., Lanier,L.L. and Somoza,C.

TITLE B70 antigen is a second ligand for CTLA-4 and CD28

JOURNAL Nature 366 (6450), 76-79 (1993)

MEDLINE 94050123

PUBMED 7694153

REFERENCE 4 (bases 1 to 478)

AUTHORS Jellis,C.L.

TITLE Direct Submission

JOURNAL Submitted (29-NOV-1994) Cindy L. Jellis, Molecular Biology, Repligen Corporation, One Kendall Square, Building 700, Cambridge, MA 02139, USA

#### FEATURES

source

Location/Qualifiers

1..478  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="DMPC-HPF no.1-1390-A1 hB7.2 genomic DNA (exon 1)"  
/sex="male"  
/cell\_type="fibroblast"  
/tissue\_type="foreskin"  
/clone\_lib="Du Pont Merck Pharmaceutical Company Human foreskin Fibroblast PI Library no.1"  
173..292  
/gene="B7.2"  
/note="alternative 5' exon"  
/citation=[2]  
/citation=[1]  
/number=1

exon

#### ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 478;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20

|||||

Db 229 CGTGTCTGTGCTAGTCCC 210

#### RESULT 6

AR534819/c

LOCUS

DEFINITION AR534819 1112 bp DNA linear PAT 08-OCT-2004

ACCESSION AR534819

SEQUENCE 25 from patent US 6734172.

VERSION AR534819.1 GI:53925465

KEYWORDS

Unknown.

SOURCE

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 1112)

AUTHORS Scholler,N.B., Disis,M.L., Hellstrom,I. and Hellstrom,K.E.

TITLE Surface receptor antigen vaccines

JOURNAL Patent: US 6734172-A 25 11-MAY-2004;

FEATURES Location/Qualifiers

source

1..1112

/organism="unknown"

/mol\_type="genomic DNA"

#### ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1112;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20

|||||

Db 57 CGTGTCTGTGCTAGTCCC 38

#### RESULT 7

HUMB72A/c

LOCUS

DEFINITION Human CTLA4 counter-receptor (B7-2) mRNA linear PRI 31-DEC-1994

ACCESSION L25259

VERSION L25259.1 GI:416368

KEYWORDS CTLA4 counter-receptor.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE 1 (bases 1 to 1112)

AUTHORS Freeman,G.J., Gribben,J.G., Boussiotis,V.A., Ng,J.W., Restivo,V.A.

Jr., Lombard,L.A., Gray,G.S. and Nadler,L.M.

TITLE Cloning of B7-2: a CTLA-4 counter-receptor that costimulates human T cell proliferation

JOURNAL Science 262 (5135), 909-911 (1993)

MEDLINE 94053735

PUBMED 7694363

COMMENT Original source text: Homo sapiens cDNA to mRNA.

FEATURES Location/Qualifiers

1..1112

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/cell\_type="B lymphocyte"

1..1112

/gene="B7-2"

107..1096

/gene="B7-2"

/codon\_start=1

/product="CTLA4 counter-receptor"

/protein\_id="AA58389.1"

/db\_xref="GI:416369"

/translation="MDPQTMGLSNILFVMAFLSGAALPKIQAYFNETADLPQCFAN  
SQNSLSLVVWQDQENLVNLYLGEKFDVSHVSGRTSPDSWTLRLNLQI  
KDGLYQCIHHKPTGMIRIHMNSLSLANFSQPEIVISNITENVINLTCSII  
HGYPEPKMSVLLRTKNSITIEDGIMOKSDQNTDLYSVISLSVSPDVTNNMTIFC  
ILETDKTRLLSSPFSIELEDPQPPPHIPMTAVLPTVILCVMYFCILIMKKKKRP  
RNSYKGTNTWREESQTKKREKIHIFERSDEAQRVFKSKTSKSDKSDTCF"

sig\_peptide

/gene="B7-2"

176..1093

/gene="B7-2"

/product="CTLA4 counter-receptor"

1112

/gene="B7-2"

polyA\_site

#### ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1112;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20

|||||

Db 57 CGTGTCTGTGCTAGTCCC 38

#### RESULT 8

AR030780/c

LOCUS

DEFINITION AR030780 1120 bp DNA linear PAT 29-SEP-1999

SEQUENCE 1 from patent US 5861310.

ACCESSION AR030780

VERSION AR030780.1 GI:5943994

KEYWORDS

Unknown.

SOURCE

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 1120)

AUTHORS Freeman,G.J., Nadler,L.M. and Gray,G.S.

TITLE Tumor cells modified to express B7-2 with increased immunogenicity

```

and uses therefor
Patent: US 5861310-A 1 19-JAN-1999;
Location/Qualifiers
1..1120
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 9
AR112747/c
LOCUS 1120 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6130316.
ACCESSION AR112747
VERSION AR112747.1 GI:14092647
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1120)
AUTHORS Freeman,G.J., Nadler,L.M., Gray,G.S. and Greenfield,E.
TITLE Fusion proteins of novel C14A4/CD28 ligands and uses therefore
JOURNAL
PATENT: US 6130316-A 1 10-OCT-2000;
Location/Qualifiers
1..1120
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 10
AR146413/c
LOCUS 1120 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 22 from patent US 6218510.
ACCESSION AR146413
VERSION AR146413.1 GI:15109602
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1120)
AUTHORS Sharpe,A.H., Borriello,F., Freeman,G.J. and Nadler,L.M.
TITLE B7-1 and B7-2 polypeptides
JOURNAL
PATENT: US 6218510-A 22 17-APR-2001;
Location/Qualifiers
1..1120
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

and uses therefor
Patent: US 5861310-A 1 19-JAN-1999;
Location/Qualifiers
1..1120
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 11
BD272169/c
LOCUS 1120 bp DNA linear PAT 17-JUL-2003
DEFINITION Use of soluble costimulatory molecules to enhance immune responses.
ACCESSION BD272169
VERSION BD272169.1 GI:33081937
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1120)
AUTHORS Sturmhoeft,K., Wolf,S.F. and OGTOOLE,M.
TITLE Use of soluble costimulatory molecules to enhance immune responses
JOURNAL
PATENT: JP 2002544170-A 2 24-DEC-2002;
GENETICS INSTITUTE INC
OS Homo sapiens (human)
PN JP 2002544170-A/2
PD 24-DEC-2002
PF 05-MAY-2000 JP 2000616813
PR 06-MAY-1999 US 60/132944
PI KNOT STURMHOFEL,STANLEY F WOLF,MARGOT O'TOOLE PC
A61K39/39,A61K39/00,A61P13/02,A61P13/08,A61P13/10,A61P15/00, PC
A61P17/00,
PC A61P35/00
CC Use of soluble costimulatory molecules to enhance immune CC
FH Key Location/Qualifiers
FT CDS (107)..(1093).
FEATURES
source
1..1120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 12
AR196804/c
LOCUS 1120 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 3 from patent US 6352694.
ACCESSION AR196804
VERSION AR196804.1 GI:20246653
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1120)
AUTHORS June,C.H., Thompson,C.B., Nabel,G.J., Gray,G.S. and Rennert,P.D.
TITLE Methods for inducing a population of T cells to proliferate using agents which recognize TCR/CD3 and ligands which stimulate an accessory molecule on the surface of the T cells
JOURNAL
PATENT: US 6352694-A 3 05-MAR-2002;
Location/Qualifiers
1..1120
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

```

```

Qy 1 CGTGTGCTCTGTGCTAGTCCC 20
Db 57 CGTGTGCTCTGTGCTAGTCCC 38

RESULT 13
LOCUS AR287728 1120 bp mRNA linear PAT 12-JUN-2003
DEFINITION Sequence 3 from patent US 6534055.
ACCESSION AR287728
VERSION AR287728.1 GI:31674734
KEYWORDS
SOURCE
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1120)
AUTHORS June, C.H., Thompson, C.B., Nabel, G.J., Gray, G.S. and Rennert, P.D.
TITLE Methods for selectively stimulating proliferation of T cells
JOURNAL Patent: US 6534055-A 3 18-MAR-2003;
FEATURES Location/Qualifiers
source 1..1120
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTCTGTGCTAGTCCC 20
Db 57 CGTGTGCTCTGTGCTAGTCCC 38

RESULT 14
LOCUS AR374141 1120 bp mRNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6605279.
ACCESSION AR374141
VERSION AR374141.1 GI:40076732
KEYWORDS
SOURCE
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1120)
AUTHORS Freeman, G.J., Nadler, L.M. and Gray, G.S.
TITLE Therapeutic compositions for inhibiting the interactions of B7-1 and B7-2 with their natural ligands
JOURNAL Patent: US 6605279-A 1 12-AUG-2003;
FEATURES Location/Qualifiers
source 1..1120
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTCTGTGCTAGTCCC 20
Db 57 CGTGTGCTCTGTGCTAGTCCC 38

RESULT 15
LOCUS AR381495 1120 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 22 from patent US 6608180.
ACCESSION AR381495
VERSION AR381495.1 GI:40089556
KEYWORDS
SOURCE
ORGANISM Unknown.

Qy 1 CGTGTGCTCTGTGCTAGTCCC 20
Db 57 CGTGTGCTCTGTGCTAGTCCC 38

RESULT 16
LOCUS AR527666 1120 bp mRNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6723705.
ACCESSION AR527666
VERSION AR527666.1 GI:53915455
KEYWORDS
SOURCE
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1120)
AUTHORS Freeman, G.J., Nadler, L.M. and Gray, G.S.
TITLE Tumor cells modified to express B7-2 with increased immunogenicity and uses therefor
JOURNAL Patent: US 6723705-A 1 20-APR-2004;
FEATURES Location/Qualifiers
source 1..1120
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTCTGTGCTAGTCCC 20
Db 57 CGTGTGCTCTGTGCTAGTCCC 38

RESULT 17
LOCUS AX047043 1120 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 3 from Patent WO0067788.
ACCESSION AX047043
VERSION AX047043.1 GI:11876461
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sturmhoefel, K., Wolf, S.F. and O'Toole, M.
TITLE Use of soluble costimulatory molecules to enhance immune responses
JOURNAL Patent: WO 0067788-A 3 16-NOV-2000;
FEATURES Location/Qualifiers
source 1..1120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
107..1096
/note="unnamed protein product"
CDS

```

```

/codon_start=1
/db_xref="GI:11876462"
/translation="MDPOCTMGLSNILFVMAFLLSGAAPLKIQAAYFNETADLPCCOFAN
SONQSLSELVWFQDOENLVNVEYLKKEPDSVHSKYMGRTSPDSKWLRLHNLQI
KDKGLYQCLIIHHKKTGMIRIHMNSELSVLANFQPEIPIPSNITENYINLTGSSI
HGYBPFGMSVLLRTKNSITIEYDGMQSDNVELYDVISLSVSPDVTNMTIFC
ILETDKRLSSPFSIELEDQPPDPHPITAVLPVTIICVMVFLILMKWKKKRRP
RNSYKCGTNTMERBESEQTKREKIHIPERSDEQVFKSKSTSSCDKSDTCF"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQTGTGTCGTGCTAGTCCC 20
Db 57 CQTGTGTCGTGCTAGTCCC 38

RESULT 18
BC040261/c
LOCUS BC040261 2781 bp mRNA linear PRI 30-JUN-2004
DEFINITION Homo sapiens CD86 antigen (CD28 antigen ligand 2, B7-2 antigen),
transcript variant 1, mRNA (CDNA clone MGC:34413 IMAGE:5173789),
complete cds.
ACCESSION BC040261 GI:25955518
VERSION BC040261.1
KEYWORDS Homo sapiens (human)
SOURCE MGC.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2781)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullah,J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2781)
Strausberg,R.
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/

Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAK Plate: 51 Row: j Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 29029570.

FEATURES
Location/Qualifiers
1..2781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:34413 IMAGE:5173789"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2781
/gene="CD86"
/note="synonyms: B7-2, B70, MGC34413, CD28LG2, LAB72"
/db_xref="LocusID:942"
/db_xref="MIM:601020"
117..1106
/gene="CD86"
/codon_start=1
/product="CD86 antigen, isoform 1"
/protein_id="AAH40261.1"
/db_xref="GI:25955519"
/db_xref="LocusID:942"
/db_xref="MIM:601020"
/translation="MDPOCTMGLSNILFVMAFLLSGAAPLKIQAAYFNETADLPCCOFAN
SONQSLSELVWFQDOENLVNVEYLKKEPDSVHSKYMGRTSPDSKWLRLHNLQI
KDKGLYQCLIIHHKKTGMIRIHMNSELSVLANFQPEIPIPSNITENYINLTGSSI
HGYBPFGMSVLLRTKNSITIEYDGMQSDNVELYDVISLSVSPDVTNMTIFC
ILETDKRLSSPFSIELEDQPPDPHPITAVLPVTIICVMVFLILMKWKKKRRP
RNSYKCGTNTMERBESEQTKREKIHIPERSDEQVFKSKSTSSCDKSDTCF"

ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 2781;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQTGTGTCGTGCTAGTCCC 20
Db 67 CQTGTGTCGTGCTAGTCCC 48

RESULT 19
AC068630
LOCUS AC068630 164161 bp DNA linear PRI 25-NOV-2002
DEFINITION Homo sapiens 3 BAC RP11-289N10 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
ACCESSION AC068630
VERSION AC068630.21 GI:16572839
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 164161)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oesman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,

```

Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Jolivet, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolly, S., Joudah, S., Karlson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisghe, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okuwonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sison, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Zorrilla, S., Zorrilla, S., Weinstein, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 164161)  
Worley, K.C.

Direct Submission  
Submitted (06-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 164161)  
Worley, K.C.

Direct Submission  
Submitted (01-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 164161)  
Worley, K.C.

Direct Submission  
Submitted (25-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 1, 2001 this sequence version replaced gi:16152225.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:  
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.  
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.  
(Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA consensus. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	source	QUALSTAT-REPORT	Location/Qualifiers
repeat_region		1. .1980	/rpt family="L1MA4"
misc_feature		complement(1. .1972)	/note="overlaps bases 28. .1999 of clone AC068754"
repeat_region		2077. .2651	/function="clone overlap"
repeat_region		2661. .3430	/rpt family="MER7C"
repeat_region		3433. .3457	/rpt family="L1ME3A"
repeat_region		complement(3463. .3984)	/rpt family="TCTCCC"
repeat_region		complement(3985. .4938)	/rpt family="SVA"
repeat_region		5066. .5492	/rpt family="L1ME3A"
repeat_region		complement(5575. .5731)	/rpt family="L1MA5"
repeat_region		5732. .6018	/rpt family="AluSp"
repeat_region		6021. .6056	/rpt family="CABA"
repeat_region		complement(6057. .6264)	/rpt family="L1MA5"
repeat_region		7148. .7175	/rpt family="AT-rich"
repeat_region		8832. .9130	/rpt family="L1MA9"
repeat_region		9183. .9383	/rpt family="AluJb"
repeat_region		9384. .9406	/rpt family="CAAAA"
repeat_region		9407. .9421	/rpt family="AluJb"
repeat_region		complement(9789. .10085)	/rpt family="L1ME3"
repeat_region		10150. .10362	/rpt family="MIR"
repeat_region		10895. .10994	/rpt family="CA-rich"
repeat_region		complement(11549. .11625)	/rpt family="L2"
repeat_region		12422. .12704	/rpt family="AluSx"
repeat_region		12705. .12745	/rpt family="CABA"
repeat_region		13095. .13253	/rpt family="MERSA"
repeat_region		complement(13567. .13768)	

```

/rpt_family="MIR"
14327..14640
/rpt_family="AluSx"
14987..15072
/rpt_family="MIR"
15187..15226
/rpt_family="(TTCA)n"
complement(15227..15332)
/rpt_family="L2"
complement(15339..15539)
/rpt_family="MIR"
complement(16577..16739)
/rpt_family="MIR"
17310..17336
/rpt_family="(GA)n"
17558..17916
/rpt_family="L2"
complement(17921..18090)
/rpt_family="MIR"
18097..18550
/rpt_family="L2"
19164..19451
/rpt_family="L2"
21617..21649
/rpt_family="GC-rich"
21812..21950
/standard_name="91222"
22117..22390
/rpt_family="L2"
22913..23020
/rpt_family="MIR"
complement(23467..23606)
/rpt_family="MIR"

Query Match      100.0%; Score 20; DB 9; Length 164161;
Best Local Similarity 100.0%; Fred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTGTCTGTCTAGTCCC 20
Db 150498 CGTGTCTGTCTAGTCCC 150517

RESULT 20
AC016918
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-23313, WORKING DRAFT SEQUENCE,
7 unordered pieces.
AC016918
VERSION AC016918.15 GI:19172803
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178804)
Muzny,D.M., Adams,C., Ali-Othman,F.R., Allen,C.,
Alsbrooks,S.L., Anarstange,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowles,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,

```

```

Homsí, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karleson, E., Kelly, S., Khan, U., King, L., Korvah, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisaeged, H.,
Lozadó, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.E., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubokan, I., Roife, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vaequez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 178804)
Worley, K.C.
Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178804)
Worley, K.C.
Direct Submission
Submitted (06-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 6, 2002 this sequence version replaced gi:17646868.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMSC
Center clone name: Rp11-23313
----- Summary Statistics
Sequencing vector: Plasmid;
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 11% of reads
Chemistry: Dye-terminator Big Dye; 89% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 207015 bases at least Q40
Consensus quality: 216378 bases at least Q30
Consensus quality: 223344 bases at least Q20
Estimated insert size: 203159; sum-of-contigs estimation
Quality coverage: 9.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2660: contig of 2660 bp in length
* 2661 2760: gap of unknown length
* 2761 5681: contig of 2921 bp in length
* 5682 5781: gap of unknown length
* 5782 8570: contig of 2789 bp in length

```

\* 8571 8670: gap of unknown length  
 \* 8671 11983: contig of 3313 bp in length  
 \* 11984 12083: gap of unknown length  
 \* 12084 17278: contig of 5195 bp in length  
 \* 17279 17378: gap of unknown length  
 \* 17379 64030: contig of 46552 bp in length  
 \* 64031 64130: gap of unknown length  
 \* 64131 178804: contig of 114674 bp in length.

## FEATURES

Location/Qualifiers  
 1..178804  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="RP11-233L3"

## ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 178804;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCCC 20  
 |||||  
 Db 9115 CGTGTCTCTGCTAGTCCC 9134

## RESULT 21

AY080604 2710 bp mRNA linear PLN 18-SEP-2002  
 Arabidopsis thaliana unknown protein (At3g05380) mRNA, partial cds.

## LOCUS

AY080604

## ACCESSION

AY080604.1 GI:19347723

## VERSION

FLJ CDNA.

## KEYWORDS

Arabidopsis thaliana (thale cress)

## SOURCE

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 2710)

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,

Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,

Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,

Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,

Davis,R.W., Ecker,J.R., and Theologis,A.

Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 2710)

Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,

Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,

Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,

Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,

Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,

Shinozaki,K., Davis,R.W., Ecker,J.R., and Theologis,A.

Direct Submission

Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN

Arabidopsis Full-length cDNA'); Seki,M., Narusaka,M., Ishida,J.,

Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,

Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, GSEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,

Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,

Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,

Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,

Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,

Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,

Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/GSEC) and Seki,M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/GSEC)  
 contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
 genome submitted to GenBank.

## FEATURES

Location/Qualifiers  
 1..2710  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="RAFL08-15-A12 (R11137)"  
 /ecotype="Columbia"  
 /note="This clone is in a modified pBluescript vector  
 (FLC-1) as a BamHI/XhoI insert."

## gene

<1..2710

## CDS

<1..2465

/gene="At3g05380"

/codon\_start=3

/evidence=experimental

/product="unknown protein"

/protein\_id="AAL86288.1"

/db\_xref="GI:19347724"

/translation="GSPKVSSEPNRRTSLSDSPISKWGMSTRKSKHCGSSIFE

EMESSRERKLDSDKDTLLMDMERAGEMAPRKRVYKRVKVEACNDSDDNGE

ACSTQGLRSQSRKAAIEASREKYSRPFKRRDDHTSGAFDPAQALAEIASMLP

ANMSELSAQLEEDTEYDMDEKSTPEATSTSHGKEANVEAPDDLLHAISVENA

NKRKSPRLVSTDCCDDPTGKLQPTSGSLRKRKPVLGDEAPAEFSQNKSNKKE

PODENNMKSLVTKRAGQVPAQSKOMKTVKALESAITSDKRPMDIVASPKQVSDS

GPTLSQKPNRRKSKLQSKAEKSETHKARSRSLSQEELLKKLATSLSP

PFARRRCIFENFSAIDHPWFSKMEFVDYLNHVGILHPIRLTRLEWSVIKSLGRPR

FSERFLHEEREKLVQYRESYKHVTELTGAREGLPTDLARPLAVGNRVIAIHPKRE

IHDGKIITVDHNKCNVLFDDGLGVLMVDIDCMPLNPIEYMEGLRRQIDKLSMKEA

LGSGNTLVGVLPFCGLNVFSFNMNPLNQGDMAPIILHGKVSSTSPRQTNHVS

ITTYNKAKEAIIORALQHALDEKEMPEMLEIVKGSKTRAQAMVDAAIKAASVKE

GEDVNTMQALELVGNKQLLRSMVYKHVNGSIEHHNPFSPNSGSEPVANNDLNS

ODGSEKQMPSELITSCVATVLMIQMCTERQYPPADVLAQLIDAVTSLQPRCPQLP

IYRIEQTCMGRIKTIQIMSLVPT"

2466..2710

/gene="At3g05380"

/misc\_difference 2694

/gene="At3g05380"

/note="not present in genomic sequence"

Query Match 87.0%; Score 17.4; DB 8; Length 2710;

Best Local Similarity 94.7%; Pred. No. 4.8e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20

|||||

Db 1726 GTGTCTCTGCTAGTCCC 1744

## RESULT 22

## LOCUS

CQ803940

## DEFINITION

Sequence 351 from Patent WO2004035798.

## ACCESSION

CQ803940

## VERSION

CQ803940.1 GI:47110604

## KEYWORDS

Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1

Inze,D., de Veyder,L. and Vlieghe,K.

Identification of novel e2f target genes and use thereof



JOURNAL Patent: WO 2004035798-A 351 29-APR-2004;  
CropDesign N.V. (BE)  
FEATURES Location/Qualifiers  
source 1. 3168  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 6; Length 3168;  
Best Local Similarity 94.7%; Pred. No. 4.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GTGTCTCTGCTAGTCCC 20  
|||||  
Db 2429 GTGTCTCTGCTATTCCT 2447

RESULT 23  
AJ583496  
LOCUS 3518 bp mRNA linear PLN 22-JUL-2004  
DEFINITION Arabidopsis thaliana mRNA for Arabidopsis always early 2 protein  
(Ataly2 gene).  
ACCESSION AJ583496  
VERSION AJ583496.1 GI:50539419  
KEYWORDS always early 2 protein; Ataly2 gene.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1  
Bhatt A.M., Zhang Q., Harris S.A., White-Cooper H. and Dickinson H.  
Gene structure and molecular analysis of Arabidopsis thaliana  
ALWAYS EARLY homologs  
Gene 336 (2), 219-229 (2004)  
PUBMED 15246533  
REFERENCE 2 (bases 1 to 3518)  
AUTHORS Bhatt A.M.  
TITLE Direct Submission  
JOURNAL Submitted (23-SEP-2003) Bhatt A.M., Department of Plant Sciences,  
University of Oxford, South Parks Road, Oxford, OX1 3RB, UNITED  
KINGDOM

## FEATURES

source Location/Qualifiers  
1. 3518  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/tissue\_type="inflorescence"  
/note="ecotype: Columbia"  
1. 3518  
/gene="aly2"  
135. 3290  
/gene="aly2"  
/codon\_start=1  
/evidence=experimental  
/product="always early 2 protein"  
/protein\_id="CAE47461.1"  
/db\_xref="GI:50539420"  
/translation="MAPVKRSVNNKFRNTSPKADGSKNNKRLKSLDGLGPGW  
TRLERFVDAVRKHGQWRVAAAIRNSRSDVMEALPNMRAVLSLPEGTASVAGL  
TAMTDDHVSVEGSGEGHDASVPRKQKRPKQSDSPREVDIQSIGSPDG  
CLIFLKQARANGTORHATGKTPRPVQTSFNRDDREGSTPNKARKQPFANDDVAH  
FLALATDARRGSGPKVSESTPNSPIKSGWMSKRSQKSGKSGSSIFEE  
WMESRRLKDDKDTLLMDMERAGMEAPKRVKRVKVEEACNDSDDNGEA  
CSATQGLRSQRKAAAEAREKYSRSPKRDHHTSGAFDALQALAEIASMLPA  
NMESLGAQLKEERTYEDMDKSTPATSTSHGEKXANVPDLSLHAISVSNAN  
KKSKPSRLVDCDDVPTGKLPQTSGLRKPVLGDEAPAFSONKSKINKELP  
QDENNMKSIVTKRAGQVPAQSKQMTVKALEESALTSKKRPGNDIVASPKQVSDSG  
PTLSQKPPNRRKSLQKLSQKSETHKAASRSLSLSEQLLLKDLKSLATSLSPF  
FARRRCIFEWFSYSAIDHPWFSKMEFVYLNHVLGHIPLRLTLEWSVKISGLGRPRF  
SERFLHEERELKQYRESVRKHYTELRTGAREGLPTDLARPLAVGNRVIAIHPKTREI

## gene

## CDS

## ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 3518;  
Best Local Similarity 94.7%; Pred. No. 4.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GTGTCTCTGCTAGTCCC 20  
|||||  
Db 2551 GTGTCTCTGCTATTCCT 2569

## RESULT 24

ATAC009606/c  
LOCUS 91924 bp DNA linear PLN 30-OCT-2002  
DEFINITION Arabidopsis thaliana chromosome III BAC F22F7 genomic sequence,  
complete sequence.  
ACCESSION AC009606  
VERSION AC009606.4 GI:12484386  
KEYWORDS HTG.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 91924)  
AUTHORS Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,  
Wu D., Maiti R., Rong C.M., Koo H., Fujii C.Y., Utterback T.R.,  
Barnstead M.E., Bowman C.L., White O., Nierman W.C. and Fraser C.M.  
Arabidopsis thaliana chromosome III BAC F22F7 genomic sequence  
Unpublished  
2 (bases 1 to 91924)  
JOURNAL Lin X. and Kaul S.  
AUTHORS Direct Submission  
TITLE Submitted (28-AUG-1999) The Institute for Genomic Research, 9712  
JOURNAL Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org  
REFERENCE 3 (bases 1 to 91924)  
AUTHORS Lin X.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-2001) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
COMMENT On Jan 25, 2001 this sequence version replaced gi:12280792.  
Address all correspondence to: atetigr.org

BAC clone F22F7 is from Arabidopsis chromosome III and is near the  
molecular marker mil72.

The orientation of the sequence is from SP6 to T7 end of the BAC  
clone.

Genes were identified by a combination of three methods: Gene  
prediction programs including GRAIL (available by anonymous ftp  
from arthur.epm.ornl.gov), Genefinder (Phil Green, University of  
Washington), Genscan (Chris Burge,  
http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene  
(http://www.cbs.dtu.dk/netpGene/cbsnetGene.html), searches of the  
complete sequence against a peptide database and the Arabidopsis  
EST database at TIGR (http://www.tigr.org/tdb/at.html).  
Annotated genes are named to indicate the level of evidence for  
their annotation. Genes with similarity to other proteins are named  
after the database hits. Genes without significant peptide  
similarity but with EST similarity are named as 'unknown' proteins.  
Genes without protein or EST similarity, that are predicted by more  
than two gene prediction programs over most of their length are  
annotated as 'hypothetical' proteins. Genes encoding tRNAs are  
predicted by tRNAscan-SE (Sean Eddy,  
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are  
identified by RepeatMasker (Arian Smit,  
http://ftp.genome.washington.edu/RM/RepeatMasker.html).  
Location/Qualifiers

## FEATURES

source	1. .91924 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="III" /map="m1172" /clone="P22P7" 1. .14310 /note="the annotation for genes within this region can be found in the overlapping bac F18C1 sequence 86291-100600" complement (14688. .>17303) /gene="P22P7.1" /note="identical to GB:AAF22525 from [Arabidopsis thaliana]" complement (join(14688. .14927,15027. .15167,15254. .15373,15468. .15555,15643. .15761,15940. .16131,16211. .16402,16508. .16615,16717. .16833,17184. .>17303)) /gene="P22P7.1" complement (join(14850. .14927,15027. .15167,15254. .15373,15468. .15555,15643. .15761,15940. .16131,16211. .16402,16508. .16615,16717. .16833,17184. .17303)) /gene="P22P7.1" /codon_start=1 /product="26S proteasome AAA-ATPase subunit RPT5a" /protein_id="AAF64530.1" /db_xref="GI:7596759" /translation="MATPMVEDTSFPEDQLASMTEDITRATRLLDNEIRILKEDAQ KTNLECSYKIKENQKIKLNQPLVGNVILEMNPEDDAEDGANIDLSQR RKGCVLQITFLPVLGVDPSLKPGLVGNKDSYLLIDTLPSYSDRVKAM EVDKPTDNDIGLEKIQELVEAIVLPMWTKERPEKLGVRPPKGVLLYGPPTGK TLMARCAAOYNATFLKAGPOLVMTFIGDGAKLVRDAFOLAKKAPACIIFIDEIDAI GTKRFDSEVSGDREVQTMLELLNQLDGFSSDERIKVIATNRADILDPALMRGRLD RKIEFPPTPEARARILQIHRSKKNVHPDYNPELARSTDDFNGAQLKAVCVAGMLA LRRDATEVHNEFDNEGIIQVQAKKASLNYYA" complement (15902. .15938) /rpt_family=" (TA)n" complement (18181. .18757) /note="ATREP6 ATREP6 ATREP6 repeat - a consensus." complement (18184. .18375) /note="ATREP9 ATREP9 ATREP9 repeat - a consensus." complement (19305. .19353) /rpt_family="POLY_A" 19434. .19505 /gene="P22P7.2" 19434. .19505 /gene="P22P7.2" /product="tRNA-Gly" /anticodons=(pos:19467. .19469,aa:Gly) complement (19554. .22286) /gene="P22P7.3" /note="identical to GB:CAA05054 from [Arabidopsis thaliana]" complement (join(19554. .19816,19914. .19949,20031. .20105,20246. .20323,20425. .20503,20596. .20757,20987. .21051,21140. .21229,21832. .21947,22115. .22286)) /gene="P22P7.3" complement (join(19709. .19816,19914. .19949,20031. .20105,20246. .20323,20425. .20503,20596. .20757,20987. .21051,21140. .21229,21832. .21947,22115. .22232)) /gene="P22P7.3" /codon_start=1 /product="alpha subunit of F-actin capping protein" /protein_id="AAF64531.1" /db_xref="GI:7596760" /translation="MADEDELLETETELSYDQKEIAKWFFLNAPAGEINYYAKDLKAV LSDEVYNEAAEAPVYVNTKTHMICLEMPSGAGDVIVSSYEINENEYLDPRTAQVAI VDHVQICITKVRPANDEELPSLYIEEYVALDAEIQRVSSYKMGSAVNCVKGKDN EGPSDELVIITAMRLSPQNCGRWSRVNIDPQESQVLDIKGLQVGAHYFEE GNVELDAKQPDSTIFQSADDAIAIANIIRHETBYLASLEVAYSKLPDNTFQDLR RKLVPVTRTLFPQNTLQSLTREVKEGLGK" complement (join (21860. .21947,22051. .22286)) /gene="P22P7.3"	CDS	complement (22107. .22232) /gene="P22P7.3" /note="unknown protein" /codon_start=1 /protein_id="AAF64545.1" /db_xref="GI:7596774" /translation="MADEDELLETETELSYDQKEIAKWFFLNAPAGEINYYAKGT" complement (<22769. .>25267) /gene="P22P7.4" /note="predicted by genemark.hmm, contains Pfam profile:PF01553 Acyltransferase" complement (join(<22769. .23281,23762. .23867,23966. .24084,24186. .24383,24484. .24549,24636. .24746,25154. .>25267)) /gene="P22P7.4" complement (join(22769. .23281,23762. .23867,23966. .24084,24186. .24383,24484. .24549,24636. .24746,25154. .25267)) /gene="P22P7.4" /note="hypothetical protein" /codon_start=1 /protein_id="AAF64532.1" /db_xref="GI:7596761" /translation="MGHFFVDKADLWKSALLFNKLDRFRIAVDDHGRATDLTAEE ESALFRMLQTVAVPLIGNACHVFMNGFNRVQVYGLEKLDALLNRPNKPLVTVSNHV ASVDDPFVIASLLPPKLDLARNLWTLCDRCFKNPVTSAFSRSVKVLPISRGEGI YQQGMDIAISKLNGGWVHIFPFGSRSRDGKTMGSARGIGRLILDADTLPMVVPFV HTGMDIMPVGASVPRIGKTVTVIIGDPIHFNDILSTEGAOHVSRKHLDAVSSRIGQ RLYDLKAQDVRYTEQSMWSHNAKTPSDRAAEI FHRVDWDSFGMGQAQFSESSPSK PIGSDDRIVRSKPRRVSPEGVSLKIKLMDSTEMGPAARGLLMMEYKSRVVSANV GRPLKAWREYFMNRGL" complement (25975. .27395) /note="similar to stress related protein GB:AAD51854 from [Vitis riparia]" complement (join(25975. .26514,26764. .26988,27262. .27395)) /gene="P22P7.5" complement (join(26041. .26514,26764. .26988,27262. .27303)) /gene="P22P7.5" /codon_start=1 /product="stress related protein, putative" /protein_id="AAF64533.1" /db_xref="GI:7596762" /translation="MATQTLAQPKLDMTKEERLKYLFQVQAAYEALLRFLIYA KAKDGLPKPGVSEAGVAVTVVGYEYKVDVPELVKMDQKMSVTELDRRVP PVKQVSAQAIISAAQIAPVARALASEVRAGVETASGMKASVYSYKTEPAKELIYAN YEPKAEQCAVSAWKKNQLPLFPRLAQVAVPTAATCFSEKYNDTVVKAEEKYRVTSYM PLVPTERTSKIPAEKAEETELRPHLD" 27792. .27922 /rpt_family=" (CAAAA)n" 27796. .27842 /rpt_family="POLY_A" complement (<29103. .29551) /gene="P22P7.6" complement (<29103. .29551) /gene="P22P7.6" complement (29103. .29462) /gene="P22P7.6" /note="unknown protein" /codon_start=1 /protein_id="AAF64534.1" /db_xref="GI:7596763" /translation="MTNTRAIYAVIATILAVISAVESTGDFGSDLDFVRAGSSSLFSG CTGSIARCIAREEEMEFSDISRRIQAQKYISYAMRRNSVPCSRGASYYNQCGA QANYSRGCSTITRCR" complement (30238. .30318) /rpt_family="POLY_A" 31392. .31423 /rpt_family="POLY_A" complement (<32605. .>35460) /gene="P22P7.7" /note="predicted by genemark.hmm" complement (join(<32605. .32726,32816. .32856,33102. .33246,33331. .33627,34033. .34104,34188. .34439,34681. .34789,34872. .34991,35356. .>35460))
misc_feature		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		gene	mrna
CDS		repeat_region	repeat_region
gene		repeat_region	repeat_region
trna		gene	mrna
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region
mrna		repeat_region	repeat_region
gene		repeat_region	repeat_region
mrna		repeat_region	repeat_region
CDS		repeat_region	repeat_region

Query Match 87.0%; Score 17.4; DB 8; Length 91924;  
 Best Local Similarity 94.7%; Pred. No. 2.4e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20  
 |||||  
 DB 76048 GTGTCTCTGCTATTCCC 76030

RESULT 25  
 AC140972/c  
 LOCUS  
 DEFINITION Felis catus clone RP86-144L3, WORKING DRAFT SEQUENCE, 11 ordered pieces.  
 ACCESSION AC140972  
 VERSION AC140972.2 GI:31126657  
 KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
 SOURCE Felis catus (cat)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 REFERENCE 1 (bases 1 to 144858)  
 AUTHORS Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K., Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-O., Legaspi,R., McDowell,J., Maduro,V.B., Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J., McGuire,C., Pearson,R.K., Portnoy,W.E., Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C., Standripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.  
 TITLE NISC Comparative Sequencing Initiative  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 144858)  
 AUTHORS Green,E.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-MAR-2003) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA  
 REFERENCE 3 (bases 1 to 144858)  
 AUTHORS Green,E.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-MAY-2003) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA  
 COMMENT On May 29, 2003 this sequence version replaced gi:28867004.  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: nisc.zoo@nhgri.nih.gov  
 ----- Project Information  
 Center project name: efa  
 Center clone name: 144L03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.  
 ----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 142769 bases at least Q40  
 Consensus quality: 143488 bases at least Q30

Consensus quality: 143725 bases at least Q20  
 Insert size: 168000; agarose-fp  
 Insert size: 143858; sum-of-contigs  
 Quality coverage: 10.45x in Q20 bases; agarose-fp  
 Quality coverage: 12.20x in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 11 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 908: contig of 908 bp in length  
 \* 909 1008: gap of unknown length  
 \* 1009 5899: contig of 4891 bp in length  
 \* 5900 5999: gap of unknown length  
 \* 6000 17014: contig of 11015 bp in length  
 \* 17015 17114: gap of unknown length  
 \* 17115 19195: contig of 2081 bp in length  
 \* 19196 19295: gap of unknown length  
 \* 19296 78277: contig of 58982 bp in length  
 \* 78278 78377: gap of unknown length  
 \* 78378 85051: contig of 6674 bp in length  
 \* 85052 87353: contig of 2202 bp in length  
 \* 87354 87453: gap of unknown length  
 \* 87454 89662: contig of 2109 bp in length  
 \* 89663 125083: contig of 35421 bp in length  
 \* 125084 125184: gap of unknown length  
 \* 125185 137715: contig of 12432 bp in length  
 \* 137716 144858: contig of 7143 bp in length.  
 ----- Location/Qualifiers  
 source  
 1..144858  
 /organism="Felis catus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9685"  
 /clone="RP86-144L3"  
 /clone\_lib="RP86"  
 misc\_feature  
 1..908  
 /note="assembly\_fragment  
 clone end:77  
 vector\_side:left"  
 misc\_feature  
 1009..5899  
 /note="assembly\_fragment"  
 misc\_feature  
 6000..17014  
 /note="assembly\_fragment"  
 misc\_feature  
 17115..19195  
 /note="assembly\_fragment"  
 misc\_feature  
 19296..78277  
 /note="assembly\_fragment"  
 misc\_feature  
 78378..85051  
 /note="assembly\_fragment"  
 misc\_feature  
 85152..87353  
 /note="assembly\_fragment"  
 misc\_feature  
 87454..89562  
 /note="assembly\_fragment"  
 misc\_feature  
 89663..125083  
 /note="assembly\_fragment"  
 misc\_feature  
 125184..137615  
 /note="assembly\_fragment"  
 misc\_feature  
 137716..144858  
 /note="assembly\_fragment  
 clone end:SP6  
 vector\_side:right"  
 -----  
 ORIGIN  
 Query Match 87.0%; Score 17.4; DB 2; Length 144858;  
 Best Local Similarity 94.7%; Pred. No. 2.2e+02;

Matches	18;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	CGTGTGTCGTCTAGTCC 19							
Db	131485	CGTGTGTCCTTGTCTAGTCC 131467							
RESULT 26									
AC136223	146492 bp DNA linear								PLN 09-AUG-2004
LOCUS	Oryza sativa (japonica cultivar-group) chromosome 5 clone								
DEFINITION	OSUNBa0093E24, complete sequence.								
ACCESSION	AC136223								
VERSION	AC136223.2 GI:42794090								
KEYWORDS	HTG.								
SOURCE	Oryza sativa (japonica cultivar-group)								
ORGANISM	Oryza sativa (japonica cultivar-group)								
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.								
	1 (bases 1 to 146492)								
AUTHORS	Chow, T.-Y., Heing, Y.-I. C., Chen, C.-S., Chen, H.-H., Liu, S.-M., Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R., Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H., Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C., Leu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y., Yu, S.-W., Wu, H.-P., Shaw, J.-F., *McCombie, W.R., *de la Bastide, M., *Spiegel, L., *Zutavern, T., *Muller, S., *Nascimben, L., *Bailja, V., *Bell, M., *Miller, B., *Katzenberger, F., *Andrade, M.V., *Dike, S., *O'Shaughnessy, A., *Palmer, L. and *Dedhia, N.								
TITLE	Oryza sativa (japonica cultivar-group) chromosome 5 clone								
JOURNAL	OSUNBa0093E24, complete sequence								
REMARK	Unpublished								
REFERENCE	*Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA								
	2 (bases 1 to 146492)								
AUTHORS	Chow, T.-Y. and Hsing, Y.-I. C.								
TITLE	Direct Submission								
JOURNAL	Submitted (31-OCT-2002) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan								
REFERENCE	3 (bases 1 to 146492)								
	Chow, T.-Y. and Hsing, Y.-I. C.								
AUTHORS	Direct Submission								
TITLE	Submitted (25-FEB-2004) Institute of Botany, Academia Sinica, 128, Section2, Academia Road, Nankang, Taipei 11529, Taiwan								
JOURNAL	4 (bases 1 to 146492)								
REFERENCE	Chow, T.-Y. and Hsing, Y.-I. C.								
AUTHORS	Direct Submission								
TITLE	Submitted (09-AUG-2004) Institute of Botany, Academia Sinica, 128, Section2, Academia Road, Nankang, Taipei 11529, Taiwan								
JOURNAL	On Feb 25, 2004 this sequence version replaced gi:24431494.								
COMMENT	This BAC clone was sequenced to phase II by the Academia Sinica and was completed to phase III by the Cold Spring Harbor Laboratory Genome Center. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. Genes were predicated from the integrated results of the following: BLASTN2.0, BLASTX2.0, GENSCAN (Chris Burge, <a href="http://genes.mit.edu/GENSCAN.html">http://genes.mit.edu/GENSCAN.html</a> ), Fgenesh ( <a href="http://www.softberry.com/">http://www.softberry.com/</a> ), Glimmerk ( <a href="http://www.softberry.com/">http://www.softberry.com/</a> ), Glimmerk ( <a href="http://www.tigr.org/softlab/glimmer/glimmer.html">http://www.tigr.org/softlab/glimmer/glimmer.html</a> ), TWINSKAN ( <a href="http://genes.cs.wustl.edu/">http://genes.cs.wustl.edu/</a> ) and GeneSplicer ( <a href="http://www.tigr.org/tdb/GeneSplicer/index.shtml">http://www.tigr.org/tdb/GeneSplicer/index.shtml</a> ). The sequence was searched against the Swiss-Prot+TrEMBL protein database, the NCBI plant EST database, the TIGR Rice Gene Index and the rice full-length cDNA database (KOME, <a href="http://cdna01.dna.affrc.go.jp/cDNA/">http://cdna01.dna.affrc.go.jp/cDNA/</a> ). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits.								

FEATURES	Location/Qualifiers	
	1..146492	
source	/organism="Oryza sativa (japonica cultivar-group)"	
	/mol_type="genomic DNA"	
gene	/cultivar="Nipponbare"	
	/db_xref="taxon:39947"	
CDS	/chromosome="5"	
	/clone="OSUNBa0093E24"	
CDS	/clone_lib="HindIII"	
	complement(603..5223)	
CDS	/gene="OSUNBa0093E24.1"	
	/note="hypothetical protein"	
CDS	complement(join(603..829,4147..4343,5204..5223))	
	/gene="OSUNBa0093E24.1"	
CDS	/codon_start=1	
	/product="hypothetical protein"	
CDS	/protein_id="AAT93982.1"	
	/db_xref="GI:51038179"	
CDS	/translation="MTDEAILFGLSVFPRNAPTVACAPLAQRHCPAGASAPPSTPS	
	RFOKRPRPVPVPPPPPPCKGISOEHLVSLRERESASDVGRCSSARAAALRGHAQG	
CDS	FGPARSGGARAEKVVPARLGISGGQARAGLGGAATTPGARLL"	
	complement(15094..16517)	
CDS	/gene="OSUNBa0093E24.2"	
	/note="hypothetical protein"	
CDS	complement(join(15094..15347,16436..16517))	
	/gene="OSUNBa0093E24.2"	
CDS	/codon_start=1	
	/product="hypothetical protein"	
CDS	/protein_id="AAT93983.1"	
	/db_xref="GI:51038180"	
CDS	/translation="MDSPTVPVGLSPPLLSILVRNESCYYTKGGGGGSGVGGGASVT	
	GDSGGAAGLLALLAHUTVVAATGGDDGDAARMESASGDGELNGGAAATRHGAR	
CDS	EMGQTKBED"	
	17588..20207	
CDS	/gene="OSUNBa0093E24.3"	
	/note="hypothetical protein"	
CDS	join(17588..18808,18897..19134,19223..19900,20020..20207)	
	/gene="OSUNBa0093E24.3"	
CDS	/codon_start=1	
	/product="hypothetical protein"	
CDS	/protein_id="AAT93984.1"	
	/db_xref="GI:51038181"	
CDS	/translation="MKRNGPAVNLKAFLEKAAKRTQBATGSPQSPQSPSCI	
	ESOMVVIYQOLENESDGSASPTTVPDVENYDIGGPNDESSDNDGGVYDIE	
CDS	HDPGLRTPISQYDVNDSDSVREYIALGCPQPKMKGDFPOHECGMRRFLPKWFFEF	
	KWLKSYVHRDAAYCFVCLYFKDSTNNHGGDAFVNGGPRNNIKSRFSKHAGVNSAHC	
CDS	EAEKYNFLWPKQTSIRSFASNSVEFKVQIARLTWSLCIRILLKQGLAFRGHDES	
	KDSNNKGNFRELQVLMFEFNKVVGNAPTQCOMIDHKIQOLIGSCAHETTKLV	
CDS	IEELHDSFALADESSDAQKQELALCLRFVNMGTQFVERFLGLVHEVDTSITLKE	
	AIKSLLLKYOLPLSKVHCFAHOLTLVAVAKENTDCAWFFGLAYLLNLVLMSCCK	
CDS	ITMLRIAGAEVMEALKGLIESQGLNQEWGLARPQDTRGSHYKTVHVMVLYVPSI	
	KLLFLKVKCEKNEBAALGAQTMLQVFGVSEFVLLHMMNEIFGYTDFCNALORREOD	
CDS	IVNALDLLEFTKAEILDVRDECGKWEKFLGTFVCKHKVAVMDMDGKIKPIQSRKF	
	FRDAINHRFADMFGLVIDRQGLNNRDEVNTELLRCWASFSAPKSFANVDNL	
CDS	VKLAFYPSDFDVEEMQLPQLNRYISDVVLVPVATAGVERVFSMMYIKNKLRSRMG	
	QEVNDCLVTFIERDFPLQVKDKDIINHFNQNK"	
CDS	27032..34322	
	/gene="OSUNBa0093E24.4"	
CDS	/note="putative polyprotein"	
	join(27032..27235,27338..29189,29499..29792,29812..30179,	
CDS	30201..30549,30856..31680,31861..32083,32210..33242,	
	33492..33575,33729..34322)	
CDS	/gene="OSUNBa0093E24.4"	
	/codon_start=1	
CDS	/product="putative polyprotein"	
	/protein_id="AAT93985.1"	
CDS	/db_xref="GI:51038182"	

```
/db_xref="UniProt/TreMBL:Q84SJS"
/translation="MTKPSKTEVEFSNVMVTLTDDFEGEDRKAMEEYINELTRBALM
RATTRTQCVIIKPGPRKPLAPVDKVDGFLNRLGSDMDDYTLKDKAYTANPPID
QTSKTDGAGTGTGTPDGRDROTAVGTOTAGTGYYPGOTGYSOAGPTAHDA
GOTSGRAGOTGTIPIQAIADPMTNASYLVHFTFDPPSTATSNQVPPHPIANAKEA
SRGYPDTRHGQYNIHAPQOTPIRPPLNPQOQGNMEDMISDVIRKFGLETNRRA
RYTKTYIDYNVFPFQYRVPFEKFSBEGDKTTWEHVQFLAQCEANSDPTFKJR
LFLSLSGNAFTWFTSLPANSIHTWAQEEQPHDYFAGETELRICDTSVVKYKNS
VIDYIKRFDVNCQYTLNITDRLDLAGLAFNGLVALLERLDGQFLDVSQMLQALRA
QBSRVKDNKRLAPDCKNNVNVDPKASDSEEGDHDIVVAEWSWTKNPKPVCNS
LMPTRKQOSKRVFQVAKCDKIDYLLQEKQJRLPKGHVIPSQELKRAYCKWH
DHSHTDNCNFRVQSAIDEGRLKFTDGSKMLDHDPPFWNTFNNDKVKLRPE
QASTKREKVIIGERSRMIIVPKPEVGVKKNRSGASTSKAPRTKTKTFDMLLEKYE
KQGEKREKRVKVEERTSEVITIKVGHSDVPIPSDGVGKSPKSEAGTSSQAGVL
TRPFGSDCGHAGATDPSGDRDPSDGTDSGRFDRVPRHVAGLIGLQRRFDRIRV
KGDVRASSQSGPMPKSTLEKAPIDIVEEKEQSDVKSESSPKEDMDINMFCAMDEAE
RDKRNELOPPVMPKSTLEKAPIDIVEEKEQSDVKSESSPKEDMDINMFCAMDEAE
VQAQSLGPKDAVFEKDESNRMKFLYLKGVHDGKRVSRMLVDGGAANVLMPSYLFKK
VAGDEDELTKTNMILNGFTEPEARGIFASBLTVGNKTLPTAFPIVDVQEANLEWLQ
KRVQYRSTKNDIGSTEDFDEKLGQGLTADPLEVDIGDGTGKPRPTFVNKNMRA
DYKVKIIELLKAGADCFAMEYHEIPGLSELVEHFRPPIKSGFRPKQPPHFNPLLYD
RVKEIDRLKAGFPCRCYAEWSSIVPEKESKIRVCIDFRLNKAATPKDVPYM
PIADMDINDASGHKVISLGDAGYNOIPEAEDMYKTAFCPCFVGLFEWVMTFGL
KNAGYQRAINLIFHLLGLVLEIYIDIVVKSDEADFTTGAKQOEAFEEELRYLST
PPVAPKAGKFRILYIASSEDKVIGAVLTQEEDEKEYIITLSRLRLDAETRPILSGR
IGKWAYALTEYDLAYEPLKSMKGOIVCDFIIDHHIDCAYEGVEVLEIVEPWKIFYNGS
SCKEGOGIQLVFPNGMCEYASARLEYCTNNQAEYNALLFGLOVMEMVGAKEYEAF
GDSLVQVQVAGVYKCLDSNLYLSDCLDIITANFDNPAIRHIAARDNSRANDLAQQA
SSVNVKRGFLJLEPGLDRLSLCEIGKIRDOGRSDRHTAGLTGNGRSOWPHATGL
TGNPERSNRCVGLTSGQNGAGHVSINLEABLIVNSDVCAQIEGDMWRIPLIQTLK
DPTLKVDRIKQAFKYTLDDGLLYRNDIDGVLKCLDGLSDGKAMGEGHRELVATD
YFTKWAEEVRLQYHTTANGAESNKTLLKLVKIKVYEPKRWHEVLSEALWAHRI
SKHGATREVPFELAYGOEAVLPEVNLGSLHVIKQDLSSEYDKMLMGNDLDEVIDDKR
LKAREIEKEKRRKRYAVAKYKAKLFOVGLDVIATLPLGTRSEKFGKWSFSGWEGY
RVCGIVRGNAVYFLEALQGERFORAINGKVLKYPFSVWQDA"
complement(38822..43272)
/gene="OSJNBa009324.5"
/translation="putative polyprotein"
complement(join(38822..39873,39931..40071,40154..40955,
41202..41582,41663..41756,41762..43272))
/gene="OSJNBa009324.5"
/codon_start=1
/product="putative polyprotein"
/protein_id="AAT93986.1"
/db_xref="GI:51038183"
/db_xref="UniProt/TreMBL:Q7XFK3"
/translation="MNNANRKRGGFNRCEFGCSIDLHLSHCPKLGKRGKREKNGEKT
NNKNNNNKSGLSQGRKLDNLRKAFQVCVAFPEPLSDVDSGGDDDKGNISDVCFM
ARGEDTVEYEDNEVAFEEAINILSAKKKYEYKQEFIIIESLKSIDRLKSLIPN
DGCSCVEYNEITKIRDVNAHDLNRRSLACNPFALHTRTLDLFLAKLLQKYQI
AFHSLFMNMI SAKLKKOPHDVLDCTGNLNMKLDKALGRVYEMEDVVKNNVLSQP
KCRKSGVMVDCENCANLEKEVSYLNNILRESEKGNLMLLDQSKVSTHNRGFGN
PYAHNSRHPVVLGVGARSCEILVXPDTNKTIVFKSGGIMSTWKSASSKSNVVHAKSPV
VASAKSSATSNVNRKTCSCFGKQGHVIGFCFLAHKOKKREIAFAKSKWQKS
GYSSKRPFRPQWLVDRPVPATERRSDQAPVPAVQKRSNRPFPVATVPREPATV
PRSDRPRFNSNLHVDICLLDLVALVARKNNWIVDSGCSRHITGDKNWFSSLK
KASKTENPEHFKTKRSKVFDSGLVNI SRYGVFKADPENPVSPVITCLVAKFDK
DVMFWHRLGHVGHDLTELSDLDIRLPLKDKDLDLEFPVSLKIRSDNGGEFK
NASFQFCNERGLEHFSRVPVQNGVVRKRVILVEMWATMLVEYKTPRPFKKI
HHFPHKTSYELRFGHQVSHLRFVCGCKFVLKSNRLDKFEARNTGLFLGYPADTH
GYRVLTKNTKIVETCEVDSFPTGCDIACITLSQVQGEDGHTFDEGDNNDHDE
VGSAGTQCAQTAGTSVRPAQVRSRPGSAGTVDADRDPPEITTSVEAAPL
HIQRHPPPEQIIGNIDVSHALTDSEWINAMKEKLENFERNTVMTLVEPPSGHNIIGTK
```

Query Match 87.0%; Score 17.4; DB 8; Length 146492;  
Best Local Similarity 94.7%; Pred. No. 2.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGTCTGTCTAGTCCC 20  
|||||

Db 104816 GTGTGTCTGTCTACTCCC 104834  
|||||

```
RESULT 27
CR391906
LOCUS
DEFINITION
Danio rerio clone DKEY-211K10, *** SEQUENCING IN PROGRESS ***, 9
unordered pieces.
ACCESSION
CR391906
VERSION
HTG; HTGS PHASE1
KEYWORDS
Danio rerio (zebrafish)
SOURCE
Danio rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
Mclay, K.
Direct Submission
Submitted (23-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 24, 2004 this sequence version replaced gi:46517964.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zk211K10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 159202 bases at least Q40
Consensus quality: 159648 bases at least Q30
Consensus quality: 160053 bases at least Q20
Insert size: 160716; sum-of-contigs
Insert size: 167493; 3.3% error; agarose-fp
Quality coverage: 6.27x in Q20 bases; sum-of-contigs Quality
coverage: 6.01x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2483: contig of 2483 bp in length
* 2484 2583: gap of 100 bp
* 2584 38899: contig of 36316 bp in length
* 38900 38999: gap of 100 bp
* 39000 55041: contig of 16042 bp in length
* 55042 55141: gap of 100 bp
* 55142 58203: contig of 3062 bp in length
* 58204 58303: gap of 100 bp
* 58304 70412: contig of 12109 bp in length
* 70413 70512: gap of 100 bp
* 70513 100144: contig of 29632 bp in length
* 100145 100244: gap of 100 bp
* 100245 110317: contig of 10073 bp in length
* 110318 110417: gap of 100 bp
* 110418 122844: contig of 12427 bp in length
* 122845 122944: gap of 100 bp
* 122945 161516: contig of 38572 bp in length.
Location/Qualifiers
1..161516
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-211K10"
/clone_lib="DanioKey"
1..2483
/note="assembly fragment:00070
fragment_chain:1"
misc_feature
```

```

misc_feature 2584..38899
              /note="assembly fragment:01225
              fragment chain:1"
misc_feature 39000..55041
              /note="assembly fragment:00464
              fragment chain:1"
misc_feature 55142..58203
              /note="assembly fragment:00087
              fragment chain:1"
misc_feature 58304..70412
              /note="assembly fragment:00114
              fragment chain:2"
misc_feature 70513..100144
              /note="assembly fragment:00640
              fragment chain:2"
misc_feature 100245..110317
              /note="assembly fragment:00215
              fragment chain:3"
misc_feature 110418..122844
              /note="assembly fragment:00326
              fragment chain:3"
misc_feature 122945..161516
              /note="assembly fragment:00891.0"

ORIGIN
Query Match      87.0%; Score 17.4; DB 2; Length 161516;
Best Local Similarity 94.7%; Pred No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20
   |||||
Db 50287 GTGTCTCTGTGCGACTGCC 50305

RESULT 28
AC129160
LOCUS
DEFINITION Rattus norvegicus clone CH230-108G14, *** SEQUENCING IN PROGRESS
ACCESSION AC129160
VERSION AC129160.3 GI:30580645
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 253259)
          Muzny,D,Marie,, Metzker,M,Lee,, Abramzon,S,, Adams,C,, Alder,J,,
          Allen,C,, Allen,H,, Alsbrooks,S,, Amin,A,, Anguiano,D,,
          Anyalebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,
          Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,,
          Biswal,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,,
          Bryant,N,, Buhay,C,, Burch,P,, Burrell,K,, Calderon,E,,
          Cardenas,V,, Carter,K,, Cavazos,I,, Ceasar,H,, Center,A,,
          Chacko,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,,
          Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Souza,L,,
          Davila,M.L., Davis,C,, Davy-Carroll,L,, De Anda,C,, Dederich,D,,
          Delgado,O., Denson,S,, Deramo,C,, Ding,Y,, Dinh,H,, Divya,K,,
          Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,,
          Egan,A,, Escotto,M,, Eugene,C,, Evans,C.A., Falls,T., Fan,G,,
          Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
          Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
          Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
          Gunaratne,P., Haaland,W., Hamil,C., Hamilton,N., Hernandez,J.,
          Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Herndon,J.,
          Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
          Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
          Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
          Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
          Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
          Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
          Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,

```

```

Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajda,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasan,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 253259)
Worley,K.C.
Direct Submission
Submitted (27-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253259)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23111082.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLEN
Center clone name: CH230-108G14
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 234406 bases at least Q40
Consensus quality: 237952 bases at least Q30
Consensus quality: 240010 bases at least Q20
Estimated insert size: 246479; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

\* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 4024: contig of 4024 bp in length  
 \* 4025 4124: gap of unknown length  
 \* 4125 248130: contig of 244006 bp in length  
 \* 248131 248230: gap of unknown length  
 \* 248231 251975: contig of 3745 bp in length  
 \* 251976 252075: gap of unknown length  
 \* 252076 253259: contig of 1184 bp in length.

## FEATURES

source  
 1. .253259  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-108G14"  
 misc\_feature  
 1. .1774  
 /note="wgs\_contig"  
 misc\_feature  
 4125. .5168  
 /note="wgs\_contig"  
 misc\_feature  
 5219. .7303  
 /note="wgs\_contig"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 253259;  
 Best Local Similarity 94.7%; Pred. No. 2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCCC 20

Db 130398 GTGTGCTGTGCTAGTCCC 130416  
 |||||

## RESULT 29

AC094647/c  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-4N1, WORKING DRAFT SEQUENCE, 7  
 unorderd pieces.

AC094647 6 GI:30466422

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

1 (bases 1 to 261391)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escotto, N., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Frazer, C. M., Gabisi, A., Gant, A., Garcia, A., Garner, T., Garza, M.,  
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,  
 Hollins, B., Howells, S., Hult, J., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowitz, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorensuhera, L., Loulaeeged, H., Lozado, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martine, E.,  
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nye, L.,  
 Nwankwelu, O., Okunolu, G., Olarnpungsoo, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,  
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Shen, H.,  
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,  
 Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Tatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Wright, D.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
 Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 261391)

Worley, K. C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 261391)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24635855.

The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GBCC

Center clone name: CH230-4N1

----- Summary Statistics

Assembly program: Atlas

Consensus quality: 249784 bases at least Q40

Consensus quality: 252189 bases at least Q30

Consensus quality: 253699 bases at least Q20

Estimated insert size: 276526; sum-of-contigs estimation

Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html)

\* NOTE: This sequence may represent more than one clone

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 7 contigs. The true order of the pieces

\* is not known and their order in this sequence record is



\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 250290: contig of 250290 bp in length  
 \* 250291: gap of unknown length  
 \* 250391 252114: contig of 1724 bp in length  
 \* 252115 252214: gap of unknown length  
 \* 252215 253404: contig of 1190 bp in length  
 \* 253405 253504: gap of unknown length  
 \* 253505 254576: contig of 1072 bp in length  
 \* 254577 254676: gap of unknown length  
 \* 254677 256299: contig of 1623 bp in length  
 \* 256300 256399: gap of unknown length  
 \* 256400 257642: contig of 1243 bp in length  
 \* 257643 257742: gap of unknown length  
 \* 257743 261391: contig of 3649 bp in length.

## FEATURES

source  
 1. 261391  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-4N1"

## misc\_feature

1. 1302  
 /note="wgs end extension  
 clone\_end:Sp6"  
 complement(7548 .8366)  
 /note="clone boundary  
 clone\_end:Sp6  
 site:EcoRI  
 end sequence:BH309507"

## misc\_feature

248018. 248389  
 /note="clone boundary  
 clone\_end:T7  
 site:EcoRI  
 end sequence:BH309505"  
 248563. 250290  
 /note="wgs end extension  
 clone\_end:T7"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 261391;  
 Best Local Similarity 94.7%; Pred. No. 2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTCTCTGTCTAGTCCC 20

Db 161920 GTGTCTCTGTCTAGTCCC 161902

## RESULT 30

AC129648/c

DEFINITION Rattus norvegicus clone CH230-8H18, WORKING DRAFT SEQUENCE, 3 linear HTG 13-NOV-2002

AC129648 unordered pieces.

VERSION AC129648.4 GI:24940664

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

SOURCE Rattus norvegicus (Norway rat)

## ORGANISM

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

## REFERENCE

1 (bases 1 to 280215)

AUTHORS Muzny,D,Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,  
 Anyalebechi,V., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,  
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
 Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,  
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,  
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,  
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,  
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
 Kwis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
 Lorenshewala,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,  
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,  
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
 Nwakoeleneh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,  
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,  
 Piazto,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,  
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Snajis,D.,  
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,  
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
 Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,  
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,V.,  
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstock,G. and Gibbs,R.A.

## Direct Submission

Unpublished

2 (bases 1 to 280215)

Worley,K.C.

## Direct Submission

Submitted (31-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 280215)

## Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

On Nov 13, 2002 this sequence version replaced gi:23096222.

The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)



```

----- Project Information
Center project name: GDWJ
Center clone name: CH230-8H18
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 237591 bases at least Q40
Consensus quality: 239484 bases at least Q30
Consensus quality: 240656 bases at least Q20
Estimated insert size: 244055; sum-of-coverage
Quality coverage: 7x in Q20 bases; sum-of-coverage
-----
* NOTE: Estimated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 255140: contig of 255140 bp in length
* 255141 255240: gap of unknown length
* 255241 274427: contig of 19187 bp in length
* 274428 274527: gap of unknown length
* 274528 280215: contig of 5688 bp in length.
*
* Location/Qualifiers
* 1..280215
*   /organism="Rattus norvegicus"
*   /mol_type="genomic DNA"
*   /db_xref="taxon:10116"
*   /clone="CH230-8H18"
*
* misc_feature
* 1..1128
*   /note="wgs contig"
*
* misc_feature
* 1179..5584
*   /note="wgs_contig"
*
ORIGIN
Query Match      87.0%; Score 17.4; DB 2; Length 280215;
Best Local Similarity 94.7%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGCTCTGTCTAGTCC 20
    |||||
Db 135089 GTGTGCTCTGTCTAGTCC 135071

RESULT 31
AC113196/c
LOCUS          AC113196          24804 bp    DNA    linear    HTG 26-FEB-2002
DEFINITION     Mus musculus clone RP23-172K11, LOW-PASS SEQUENCE SAMPLING.
ACCESSION      AC113196
VERSION        AC113196.1 GI:18921382
KEYWORDS       HTG; HGSC PHASE0.
SOURCE         Mus musculus (house mouse)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 24804)
AUTHORS        Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE          Mus musculus, clone RP23-172K11
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 24804)
AUTHORS        Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
               Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
               Brown,A., Camarata,J., Campopiano,A., Chang,J.J., Chazaro,B.,
               Choepel,Y., Collangelo,M., Collins,S., Collamore,A., Cook,A.,
               Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
               Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
               Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
               Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
               Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
               Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
               Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
               McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
               Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
               Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
               Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
               Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
               Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
               Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
               Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
               Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
               Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
               Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
               Direct Submission
Submitted (26-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu
----- Project Information
Center project name: L23475
Center clone name: 172_K_11
-----
* NOTE: This record contains 31 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 716: contig of 716 bp in length
* 717 816: gap of 100 bp
* 817 1520: contig of 704 bp in length
* 1521 1620: gap of 100 bp
* 1621 2327: contig of 707 bp in length
* 2328 2427: gap of 100 bp
* 2428 3149: contig of 722 bp in length
* 3150 3249: gap of 100 bp
* 3250 3968: contig of 719 bp in length
* 3969 4068: gap of 100 bp
* 4069 4762: contig of 694 bp in length
* 4763 4862: gap of 100 bp
* 4863 5567: contig of 705 bp in length
* 5568 5667: gap of 100 bp
* 5668 6345: contig of 678 bp in length
* 6346 6446: gap of 100 bp
* 6446 7147: contig of 702 bp in length
* 7148 7247: gap of 100 bp
* 7248 7966: contig of 719 bp in length
* 7967 8066: gap of 100 bp
* 8067 8741: contig of 675 bp in length
* 8742 8841: gap of 100 bp
* 8842 9561: contig of 720 bp in length
* 9562 9661: gap of 100 bp
* 9662 10355: contig of 694 bp in length
* 10356 10455: gap of 100 bp
* 10456 11371: contig of 682 bp in length
* 11372 11377: gap of 100 bp
* 11378 11931: contig of 694 bp in length
* 11932 12031: gap of 100 bp
* 12032 12747: contig of 716 bp in length
* 12748 12847: gap of 100 bp
* 12848 13532: contig of 685 bp in length
* 13533 13632: gap of 100 bp
* 13633 14312: contig of 680 bp in length

```

TITLE  
JOURNAL

## COMMENT

\* 14313 14412: gap of 100 bp  
\* 14413 15104: contig of 692 bp in length  
\* 15105 15204: gap of 100 bp  
\* 15205 15922: contig of 718 bp in length  
\* 15923 16022: gap of 100 bp  
\* 16023 16744: contig of 722 bp in length  
\* 16745 16844: gap of 100 bp  
\* 16845 17550: contig of 706 bp in length  
\* 17551 17650: gap of 100 bp  
\* 17651 18354: contig of 704 bp in length  
\* 18355 18454: gap of 100 bp  
\* 18455 19178: contig of 724 bp in length  
\* 19179 19278: gap of 100 bp  
\* 19279 19589: contig of 711 bp in length  
\* 19590 20089: gap of 100 bp  
\* 20090 20807: contig of 718 bp in length  
\* 20808 20907: gap of 100 bp  
\* 20908 21601: contig of 694 bp in length  
\* 21602 21701: gap of 100 bp  
\* 21702 22396: contig of 695 bp in length  
\* 22397 22496: gap of 100 bp  
\* 22497 23210: contig of 714 bp in length  
\* 23211 23310: gap of 100 bp  
\* 23311 24003: contig of 693 bp in length  
\* 24004 24103: gap of 100 bp  
\* 24104 24804: contig of 701 bp in length.

FEATURES

source  
1..24804  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-172K11"  
/clone\_lib="RPCL-23 Female Mouse BAC"

ORIGIN

Query Match 85.0%; Score 17; DB 2; Length 24804;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GTGTCTGTGCTAGTCCC 20  
|||||

Db 4721 GTGTCTGTGCTAGTCCC 4705

RESULT 32

LMFLCHR15\_6

WPCOMMENT

Sequence split into 8 fragments LOCUS LMFLCHR15 Accession AL160371

Fragment Name	Begin	End
LMFLCHR15_0	1	110000
LMFLCHR15_1	100001	210000
LMFLCHR15_2	200001	310000
LMFLCHR15_3	300001	410000
LMFLCHR15_4	400001	510000
LMFLCHR15_5	500001	610000
LMFLCHR15_6	600001	710000
LMFLCHR15_7	700001	720704

Continuation (7 of 8) of LMFLCHR15 from base 600001 (AL160371 Leishmania major chromosome)

Query Match 85.0%; Score 17; DB 2; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGTGTCTGTGCTAGTCCC 19  
|||||

Db 52209 TGTGTCTGTGCTAGTCCC 52225

RESULT 33

CR382128\_04/c

WPCOMMENT

Sequence split into 31 fragments LOCUS CR382128 Accession CR382128

Fragment Name	Begin	End
---------------	-------	-----

CR382128_00	1	110000
CR382128_01	100001	210000
CR382128_02	200001	310000
CR382128_03	300001	410000
CR382128_04	400001	510000
CR382128_05	500001	610000
CR382128_06	600001	710000
CR382128_07	700001	810000
CR382128_08	800001	910000
CR382128_09	900001	1010000
CR382128_10	1000001	1110000
CR382128_11	1100001	1210000
CR382128_12	1200001	1310000
CR382128_13	1300001	1410000
CR382128_14	1400001	1510000
CR382128_15	1500001	1610000
CR382128_16	1600001	1710000
CR382128_17	1700001	1810000
CR382128_18	1800001	1910000
CR382128_19	1900001	2010000
CR382128_20	2000001	2110000
CR382128_21	2100001	2210000
CR382128_22	2200001	2310000
CR382128_23	2300001	2410000
CR382128_24	2400001	2510000
CR382128_25	2500001	2610000
CR382128_26	2600001	2710000
CR382128_27	2700001	2810000
CR382128_28	2800001	2910000
CR382128_29	2900001	3010000
CR382128_30	3000001	3066374

Continuation (5 of 31) of CR382128 from base 400001 (CR382128 Yarrowia lipolytica chromosome)

Query Match 85.0%; Score 17; DB 8; Length 110000;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGTGTCTGTGCTAGTC 18  
|||||

Db 91229 GTGTGTCTGTGCTAGTC 91213

RESULT 34

BC076559/c

LOCUS

DEFINITION

Danio rerio zgc:92474, mRNA (cdna clone MGC:92474 IMAGE:7044589),  
complete cds.

ACCESSION BC076559

VERSION BC076559.1 GI:49901363

KEYWORDS MGC.

SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

REFERENCE

AUTHORS

1 (bases 1 to 2575)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinska, M.I., Skalska, U., Smal, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932  
2 (bases 1 to 2575)  
Strausberg, R.  
Direct Submission  
Submitted (06-JUL-2004) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Len Zon, Harvard  
CDNA Library Preparation: Open Biosystems  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgeon, C., Vogt, J.B., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 178 Row: p Column: 17  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein.

FEATURES  
source  
1..2575  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="MGC:92474 IMAGE:7044589"  
/tissue\_type="Whole, adult zebrafish"  
/clone\_lib="NIH ZGC\_10"  
/lab\_host="DH10B"  
/note="Vector: pExpress1"  
1..2575  
/gene="zgc:92474"  
/db\_xref="LocusID:436917"  
156..575  
/gene="zgc:92474"  
/codon\_start=1  
/product="zgc:92474"  
/protein\_id="AAH76559.1"  
/db\_xref="GI:49901364"  
/db\_xref="LocusID:436917"  
/translation="MAQVKSIAQSLYPCSQRLHQMQLIKNSVTCNDGTGAGYY  
IKESGRRLVFLGGWYCFSKHYCDRSYEMRLMSSNPRTGTGILSPQPEE  
NPHWMNANTVFYPCSSDVSWSGSTPKTDQRHQGALN"

Query Match 84.0%; Score 16.8; DB 5; Length 2575;  
Best Local Similarity 90.0%; Pred. No. 9.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGTGTCTGTGCTAGTCCC 20  
|||||  
Db 101 CCGTGTGTGTGTCTGTCTCC 82  
|||||

RESULT 35  
Query Match 84.0%; Score 16.8; DB 5; Length 2575;  
Best Local Similarity 90.0%; Pred. No. 9.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGTGTCTGTGCTAGTCCC 20  
|||||  
Db 101 CCGTGTGTGTGTCTGTCTCC 82  
|||||

RESULT 36  
Query Match 84.0%; Score 16.8; DB 9; Length 14235;  
Best Local Similarity 90.0%; Pred. No. 7e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGTGTCTGTGCTAGTCCC 20  
|||||  
Db 5757 CCGTGTCTGTGCTATTCCTCC 5738  
|||||

RESULT 36  
AC118055  
LOCUS  
DEFINITION  
Homo sapiens chromosome 15 clone CTD-2013P3 map 15, LOW-PASS  
SEQUENCE SAMPLING.  
AC118055  
VERSION  
AC118055.1 GI:20143784  
KEYWORDS  
HTG; HTGS\_PHASE0.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 58969)  
AUTHORS  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE  
Homo sapiens chromosome 15, clone CTD-2013P3  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 58969)  
AUTHORS  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collumore, A.,  
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

AF220172S2/c  
LOCUS  
DEFINITION  
Homo sapiens acid ceramidase (ASAH) gene, exons 2 through 4.  
ACCESSION  
AF220173 AF220174  
VERSION  
AF220173.1 GI:9651700  
KEYWORDS  
2 of 3  
SEGMENT  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 14235)  
AUTHORS  
Mukherjee, A.B.  
TITLE  
Human acid ceramidase gene  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 14235)  
AUTHORS  
Mukherjee, A.B.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (29-DEC-1999) HDB, NICHD/NIH, 9000 Rockville Pike,  
Bethesda, MD 20892-1830, USA  
FEATURES  
Location/Qualifiers  
1..14235  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22-21.2"  
6302..6348  
/gene="ASAH"  
/number=2  
10518..10608  
/gene="ASAH"  
/number=3  
12035..12121  
/gene="ASAH"  
/number=4  
exon  
exon  
exon  
ORIGIN

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, I.,  
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thumann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
 Submitted (12-APR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submission@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L26226  
 Center clone name: 2013\_P\_3  
 -----

\* NOTE: This record contains 74 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1  
 698 697: contig of 697 bp in length  
 797: gap of 100 bp  
 1493: contig of 696 bp in length  
 1593: gap of 100 bp  
 1594 2271: contig of 678 bp in length  
 2272 2371: gap of 100 bp  
 2372 3081: contig of 710 bp in length  
 3082 3181: gap of 100 bp  
 3182 3909: contig of 728 bp in length  
 3910 4009: gap of 100 bp  
 4010 4683: contig of 674 bp in length  
 4684 4783: gap of 100 bp  
 4784 5499: contig of 716 bp in length  
 5500 5599: gap of 100 bp  
 5600 6287: contig of 688 bp in length  
 6288 6387: gap of 100 bp  
 6388 7081: contig of 694 bp in length  
 7082 7181: gap of 100 bp  
 7182 7882: contig of 701 bp in length  
 7883 7982: gap of 100 bp  
 7983 8683: contig of 701 bp in length  
 8684 8783: gap of 100 bp  
 8784 9484: contig of 701 bp in length  
 9485 9584: gap of 100 bp  
 9585 10293: contig of 709 bp in length  
 10294 10393: gap of 100 bp  
 10394 11087: contig of 694 bp in length  
 11088 11187: gap of 100 bp  
 11188 11907: contig of 720 bp in length  
 11908 12007: gap of 100 bp  
 12008 12592: contig of 685 bp in length  
 12593 12792: gap of 100 bp  
 12793 13489: contig of 697 bp in length  
 13490 13589: gap of 100 bp

13590 14301: contig of 712 bp in length  
 14302 14401: gap of 100 bp  
 14402 15112: contig of 711 bp in length  
 15113 15212: gap of 100 bp  
 15213 15896: contig of 684 bp in length  
 15897 15996: gap of 100 bp  
 15997 16686: contig of 690 bp in length  
 16687 16786: gap of 100 bp  
 16787 17491: contig of 705 bp in length  
 17492 17591: gap of 100 bp  
 17592 18292: contig of 701 bp in length  
 18293 18392: gap of 100 bp  
 18393 19107: contig of 715 bp in length  
 19108 19207: gap of 100 bp  
 19208 19927: contig of 720 bp in length  
 19928 20027: gap of 100 bp  
 20028 20720: contig of 693 bp in length  
 20721 20820: gap of 100 bp  
 20821 21523: contig of 703 bp in length  
 21524 21623: gap of 100 bp  
 21624 22335: contig of 712 bp in length  
 22336 22436: gap of 100 bp  
 22437 23116: contig of 681 bp in length  
 23117 23216: gap of 100 bp  
 23217 23911: contig of 695 bp in length  
 23912 24011: gap of 100 bp  
 24012 24706: contig of 695 bp in length  
 24707 24806: gap of 100 bp  
 24807 25484: contig of 678 bp in length  
 25485 25584: gap of 100 bp  
 25585 26268: contig of 684 bp in length  
 26269 26368: gap of 100 bp  
 26370 27078: contig of 711 bp in length  
 27079 27178: gap of 100 bp  
 27179 27865: contig of 686 bp in length  
 27866 27965: gap of 100 bp  
 27966 28667: contig of 702 bp in length  
 28668 28767: gap of 100 bp  
 28768 29457: contig of 690 bp in length  
 29458 29557: gap of 100 bp  
 29559 30259: contig of 702 bp in length  
 30260 30359: gap of 100 bp  
 30360 31028: contig of 669 bp in length  
 31029 31128: gap of 100 bp  
 31129 31827: contig of 699 bp in length  
 31828 31927: gap of 100 bp  
 31928 32647: contig of 720 bp in length  
 32648 32747: gap of 100 bp  
 32748 33458: contig of 711 bp in length  
 33459 33558: gap of 100 bp  
 33559 34257: contig of 699 bp in length  
 34258 34357: gap of 100 bp  
 34358 35045: contig of 688 bp in length  
 35046 35145: gap of 100 bp  
 35146 35869: contig of 724 bp in length  
 35870 35969: gap of 100 bp  
 35970 36652: contig of 683 bp in length  
 36653 36752: gap of 100 bp  
 36753 37449: contig of 697 bp in length  
 37450 37549: gap of 100 bp  
 37550 38233: contig of 684 bp in length  
 38234 38333: gap of 100 bp  
 38334 39050: contig of 717 bp in length  
 39051 39150: gap of 100 bp  
 39151 39831: contig of 681 bp in length  
 39832 39931: gap of 100 bp  
 39932 40610: contig of 679 bp in length  
 40611 40710: gap of 100 bp  
 40711 41415: contig of 705 bp in length  
 41416 41515: gap of 100 bp  
 41516 42193: contig of 678 bp in length  
 42194 42293: gap of 100 bp  
 42294 42984: contig of 691 bp in length

TITLE  
JOURNAL

COMMENT



\* 16787 16786: gap of 100 bp bp in length  
\* 16787 17491: contig of 705 bp in length  
\* 17492 17591: gap of 100 bp  
\* 17592 18292: contig of 701 bp in length  
\* 18293 18392: gap of 100 bp  
\* 18393 19107: contig of 715 bp in length  
\* 19108 19207: gap of 100 bp  
\* 19208 19207: contig of 720 bp in length  
\* 19208 20027: gap of 100 bp  
\* 20028 20720: contig of 693 bp in length  
\* 20721 20820: gap of 100 bp  
\* 20821 21523: contig of 703 bp in length  
\* 21524 21623: gap of 100 bp  
\* 21624 22335: contig of 712 bp in length  
\* 22336 22435: gap of 100 bp  
\* 22436 23116: contig of 681 bp in length  
\* 23117 23216: gap of 100 bp  
\* 23217 23911: contig of 695 bp in length  
\* 23912 24011: gap of 100 bp  
\* 24012 24706: contig of 695 bp in length  
\* 24707 24806: gap of 100 bp  
\* 24807 25484: contig of 678 bp in length  
\* 25485 25584: gap of 100 bp  
\* 25585 26268: contig of 684 bp in length  
\* 26269 26368: gap of 100 bp  
\* 26369 27079: contig of 711 bp in length  
\* 27080 27179: gap of 100 bp  
\* 27180 27865: contig of 686 bp in length  
\* 27866 27965: gap of 100 bp  
\* 27966 28667: contig of 702 bp in length  
\* 28668 28767: gap of 100 bp  
\* 28768 29457: contig of 690 bp in length  
\* 29458 29557: gap of 100 bp  
\* 29558 30259: contig of 702 bp in length  
\* 30260 30359: gap of 100 bp  
\* 30360 31028: contig of 669 bp in length  
\* 31029 31128: gap of 100 bp  
\* 31129 31827: contig of 699 bp in length  
\* 31828 31928: gap of 100 bp  
\* 31928 32847: contig of 720 bp in length  
\* 32848 32747: gap of 100 bp  
\* 32748 33458: contig of 711 bp in length  
\* 33459 33558: gap of 100 bp  
\* 33559 34257: contig of 699 bp in length  
\* 34258 34357: gap of 100 bp  
\* 34358 35045: contig of 688 bp in length  
\* 35046 35145: gap of 100 bp  
\* 35146 35869: contig of 724 bp in length  
\* 35870 35969: gap of 100 bp  
\* 35970 36522: contig of 683 bp in length  
\* 36553 36752: gap of 100 bp  
\* 36753 37449: contig of 697 bp in length  
\* 37450 37549: gap of 100 bp  
\* 37550 38233: contig of 684 bp in length  
\* 38234 38333: gap of 100 bp  
\* 38334 39050: contig of 717 bp in length  
\* 39051 39150: gap of 100 bp  
\* 39151 39831: contig of 681 bp in length  
\* 39832 39931: gap of 100 bp  
\* 39932 40610: contig of 679 bp in length  
\* 40611 40710: gap of 100 bp  
\* 40711 41415: contig of 705 bp in length  
\* 41416 42193: gap of 100 bp  
\* 42194 42293: gap of 100 bp  
\* 42294 42984: contig of 691 bp in length  
\* 42985 43085: gap of 100 bp  
\* 43085 43776: contig of 692 bp in length  
\* 43777 43876: gap of 100 bp  
\* 43877 44582: contig of 706 bp in length  
\* 44583 44682: gap of 100 bp  
\* 44683 45387: contig of 705 bp in length  
\* 45388 45487: gap of 100 bp

\* 45488 46202: contig of 715 bp in length  
\* 46203 46302: gap of 100 bp  
\* 46303 47018: contig of 716 bp in length  
\* 47019 47118: gap of 100 bp  
\* 47119 47822: contig of 704 bp in length  
\* 47823 47922: gap of 100 bp  
\* 47923 48605: contig of 683 bp in length  
\* 48606 48705: gap of 100 bp  
\* 48706 49416: contig of 711 bp in length  
\* 49417 49516: gap of 100 bp  
\* 49517 50207: contig of 691 bp in length  
\* 50208 50307: gap of 100 bp  
\* 50308 50989: contig of 682 bp in length  
\* 50990 51089: gap of 100 bp  
\* 51090 51781: contig of 692 bp in length  
\* 51782 51881: gap of 100 bp  
\* 51882 52575: contig of 694 bp in length  
\* 52576 52675: gap of 100 bp  
\* 52676 53385: contig of 710 bp in length  
\* 53386 53485: gap of 100 bp  
\* 53486 54177: contig of 692 bp in length  
\* 54178 54277: gap of 100 bp  
\* 54278 54964: contig of 687 bp in length

Query Match 84.0%; Score 16.8; DB 2; Length 58969;  
Best Local Similarity 90.0%; Pred. No. 5.3e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20  
|||||  
Db 40417 CGTGTCTGTGCAAGTCTC 40398

RESULT 38  
AC004252/c

LOCUS Drosophila melanogaster (P1 DS09121 (D88)) DNA linear INV 01-MAR-1998  
DEFINITION AC004252 Drosophila melanogaster (P1 DS09121 (D88)) DNA sequence, complete  
sequence.

AC004252 AC001962 AC001964 AC001957 AC002935 AC001961 AC001959  
AC001421 AC001954 AC001417 AC001420 AC001955 AC001418 AC001958  
AC001416 AC001960 AC001953 AC001952 AC001956 AC001963  
AC004252.1 GI:2920806  
HTG.

ORGANISM Drosophila melanogaster (fruit fly)

KEYWORDS Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 59498)

Arcaini, T.T., Baxter, E., Blazej, R.G.,  
Chew, M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A.,  
Humasti, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B.,  
Lomotan, M.A., Mazda, P., Mok, M.S., Moshrefi, A.R., Moshrefi, M.,  
Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Punch, D., Snir, E.,  
Twomey, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.L.  
and Kimmel, B.

Sequencing of Drosophila chromosome 2R, region 58D1-58F1  
Unpublished (1997)

2 (bases 1 to 59498)

Chew, M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A.,  
Humasti, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B.,  
Lomotan, M.A., Mazda, P., Mok, M.S., Moshrefi, A.R., Moshrefi, M.,  
Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Punch, D., Snir, E.,  
Twomey, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.L.  
and Kimmel, B.

Direct Submission

Submitted (28-FEB-1998) Berkeley Drosophila Genome Project, MS  
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, US

Sequence submitted by:

Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our sequence  
 archive web site  
 (http://fruitfly.berkeley.edu/sequence-archive.html) or send email  
 to drosophila.genome@lbl.gov.

Library location: 1\_96.  
 This P1 was assembled from the following subclones: 2\_g7  
 (AC001962), 2\_h4 (AC001964), 2\_d6 (AC001957), 2\_h12 (AC002935),  
 2\_a3, 2\_g12 (AC001961), 2\_e3 (AC001959), 2\_g6 (AC001421), 1\_g10  
 (AC001954), 1\_g4 (AC001417), 2\_d8 (AC001420), 1\_h7 (AC001955), 2\_b2  
 (AC001418), 2\_e10 (AC001958), 1\_d2 (AC001416), 2\_b3 (AC001419),  
 2\_f1 (AC001960), 1\_c3 (AC001953), 1\_a3 (AC001952), 2\_all  
 (AC001956), 2\_g9 (AC001963).

## FEATURES

source

1..59498  
 Location/Qualifiers  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /chromosomes="2R"  
 /map="58D1-58F1"  
 /clone="P1 DS09121 (D88)"  
 /note="Does not seem to hit D62, hits D61 on sp6 end, NOT  
 T7 end which is indicated be the map!"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 3; Length 59498;  
 Best Local Similarity 90.0%; Pred. No. 5.3e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCC 20  
 |||||  
 DB 57708 CGTGTCTCTGCTGTGCC 57689

## RESULT 39

AC017870

LOCUS 88933 bp DNA linear HTG 09-DEC-1999  
 DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*.

AC017870

AC017870.1 GI:6553320

HTG; HTGS\_PHASE2.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 88933)

Adams, M. and Venter, J.C.

Direct Submission

Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,

Rockville, MD, USA

This sequence was identified as CDM:10212423 by the submitter.

For more information on this record e-mail to fly@celera.com.

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

## FEATURES

source

1..88933  
 Location/Qualifiers  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 88933;  
 Best Local Similarity 90.0%; Pred. No. 4.9e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCC 20  
 |||||  
 DB 74708 CGTGTCTCTGCTGTGCC 74727

## RESULT 40

AC151836 3/c

WPCOMMENT

Sequence split into 5 fragments LOCUS AC151836 Accession AC151836

Fragment Name	Begin	End
AC151836.0	1	110000
AC151836.1	100001	210000
AC151836.2	200001	310000
AC151836.3	300001	410000
AC151836.4	400001	485063

Continuation (4 of 5) of AC151836 from base 300001 (AC151836 Mus musculus chromosome 14 c

## Query Match

84.0%; Score 16.8; DB 2; Length 110000;  
 Best Local Similarity 90.0%; Pred. No. 4.7e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

## QY

1 CGTGTCTCTGCTAGTCC 20  
 |||||  
 DB 28923 CGTGTCTCTGCTAGTCC 28904

## RESULT 41

AC151419/c

LOCUS

DEFINITION AC151419 113386 bp DNA linear HTG 04-SEP-2004  
 Bos taurus clone CH240-13F11, WORKING DRAFT SEQUENCE, 12 unordered  
 pieces.

AC151419

AC151419.2 GI:51889783

HTG; HTGS\_PHASE1; HTGS\_DRAFT.

KEYWORDS

SOURCE

Bos taurus

Bos taurus

## ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovinae; Bos.

1 (bases 1 to 113386)

Muzny, D., Metzker, M., Adams, C., Agbai II, O., Allen, C.,

Alberbrooks, S., Archer, P., Arredondo, H., Bandaranaike, D., Bangura, L.,

Beltratti, B., Beltratti, R., Beraducci, A., Biswal, K., Blyth, P.,

Bonham, H., Buhay, C., Burch, P., Cadoree, I., Canada, A., Cardenas, V.,

Carter, K., Cavazos, I., Chacko, J., Chahrour, M., Chavez, D., Chen, A.,

Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R.,

Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M. L., Davis, C.,

Davy-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C.,

Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A.,

Durbin, K., Diu, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C.,

Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,

Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,

Gaekin, C., Gensch, S., Ghose, S., Gill, R., Gonzalez, D.,

Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K.,

Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,

Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,

Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B.,

Howell, L. T., Hulyk, S., Hume, J., Iino, K., Jackson, A., Jackson, L.,

Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,

Keys, T., Khan, Z., King, L., Kovar, C., Kowis, C., Kowis, C., Lara, F.,

Leal, S., Lee, K., Lee, S., LeGall, F. I., Lemon, S., Lewis, L., Li, B.,

Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,

London, P., Lopez, J., Lorenz, L., Lozano, R., Luk, T., Madu, R.,

Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,

McCllelland, H., McPherson, J., Mercadao, C., Milosavljevic, A.,

Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D.,

Nazarith, L., Ngo, D., Nguyen, N., Norwig-Eastough, E., Nott, A.,

Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odoh, E., Okwuonu, G.,

Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H.,

Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L.,

Puzos, M., Qin, X., Quinn, A., Quinlan, J., Rabata, D., Rachlin, E.,

Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F.,

Rojas, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R.,

Savery, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A.,

Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, B.,

Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Usmani, K.,

Vargo, C., Verduzco, D., Villasana, D., Virk, D., Volkov, A.,

Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X.,  
Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J.,  
Yakub, S., Yan, K., Yaun, Y., Yu, P., Zhang, J., Zhang, L., Zhang, Z.,  
Zhou, J., Weinstock, G. and Gibbs, R.

## Direct Submission

Unpublished  
2 (bases 1 to 113386)

Worley, K.C.

## Direct Submission

Submitted (27-AUG-2004) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 113386)

Worley, K.C.

## Direct Submission

Submitted (04-SEP-2004) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Sep 4, 2004 this sequence version replaced gi:51571973.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help.tmc.edu

----- Project Information

Center project name: FAOJ

Center clone name: CH240-13F11

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 110313 bases at least Q40

Consensus quality: 110834 bases at least Q30

Consensus quality: 111456 bases at least Q20

Estimated insert size: 119536; sum-of-contigs estimation

Estimated insert size: 116681; agarose-fp estimation

Quality coverage: 3x in Q20 bases; agarose-fp estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently  
consists of 12 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence.

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 6636: contig of 6636 bp in length

\* 6637 6736: gap of unknown length

\* 6737 10947: contig of 4211 bp in length

\* 10948 11047: gap of unknown length

\* 11048 28138: contig of 17091 bp in length

\* 28139 28238: gap of unknown length

\* 28239 32060: contig of 3822 bp in length

\* 32061 32160: gap of unknown length

\* 32161 47757: contig of 15597 bp in length

\* 47758 47857: gap of unknown length

\* 47858 52314: contig of 4457 bp in length

\* 52315 52414: gap of unknown length

\* 52415 79094: contig of 26680 bp in length

\* 79095 79194: gap of unknown length

\* 79195 83988: contig of 4794 bp in length

\* 83989 84088: gap of unknown length

\* 84089 91962: contig of 7874 bp in length

\* 91963 92062: gap of unknown length

\* 92063 100259: contig of 8197 bp in length

\* 100260 100359: gap of unknown length

\* 100360 105935: contig of 5576 bp in length

\* 105936 106035: gap of unknown length

\* 106036 113386: contig of 7351 bp in length.

Location/Qualifiers

## source

1. 113386  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9913"  
/clone="CH240-13F11"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 113386;  
Best Local Similarity 90.0%; Pred. No. 4.7e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20

|||||

Db 56342 CGTGTGCTGTGCTAGTCCC 56323

## RESULT 42

AL669954

LOCUS

DEFINITION

AL669954

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Submitted (18-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquary@sanger.ac.uk  
On Jun 21, 2002 this sequence version replaced gi:2121353.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em, EMBL; Sw,  
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP  
database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-456018 is  
from the RPCI-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>

## FEATURES

## source

1. 113279  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone="RP23-456018"  
/clone\_lib="RPCI-23"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 113279;  
Best Local Similarity 90.0%; Pred. No. 4.6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20

|||||

## FEATURES



Db 86209 COTGTCTGTCTAGTCC 86228

RESULT 43  
AC141138/c

LOCUS AC141138 126056 bp DNA linear HTG 09-MAY-2003  
DEFINITION Rattus norvegicus clone CH230-178N14, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 60 unordered pieces.

AC141138.3 GI:30466903  
AC141138 HTG: HTGS PHASE1  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 126056)  
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, Y., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoggess, M.,  
Hollins, B., Howells, S., Hui, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenshewa, L., Loulsegged, H., Lozano, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,  
Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,  
Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,  
Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G.,  
Olanrunso, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,  
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,  
Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,  
Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,  
Reuter, M., Richards, S., Riggs, R., Rives, C., Rodkey, T., Rojas, A.,  
Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S.,  
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,  
Sisson, I., Sitter, C. D., Smales, D., Sneed, A., Sodergren, E.,  
Song, X.-Z., Sorelle, R., Soia, J., Steimle, M., Strong, R., Sutton, A.,  
Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, A.,  
Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villasana, D.,  
Walton, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,  
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R.,  
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,  
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,  
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,  
Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

Direct Submission  
Unpublished

2 (bases 1 to 126056)  
Worley, K. C.

REFERENCE  
JOURNAL  
AUTHORS  
TITLE

Submitted (09-MAR-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 126056)  
Worley, K. C.  
Direct Submission  
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 9, 2003 this sequence version replaced gi:28913269.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: KDJT  
Center clone name: CH230-178N14  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 112572 bases at least Q40  
Consensus quality: 119554 bases at least Q30  
Consensus quality: 124684 bases at least Q20  
Estimated insert size: 104198; sum-of-contigs estimation  
Quality coverage: 2x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 60 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1055: contig of 1055 bp in length  
1056 1155: gap of unknown length  
1156 2369: contig of 1214 bp in length  
2370 2469: gap of unknown length  
2470 3476: contig of 1007 bp in length  
3477 3576: gap of unknown length  
3577 4932: contig of 1356 bp in length  
4933 5032: gap of unknown length  
5033 6748: contig of 1716 bp in length  
6749 6848: gap of unknown length  
6849 8053: contig of 1205 bp in length  
8054 8153: gap of unknown length  
8154 9614: contig of 1461 bp in length  
9615 9715: contig of 1212 bp in length  
9715 10926: gap of unknown length  
10927 11026: contig of 1434 bp in length  
11027 12460: contig of 1468 bp in length  
12461 14028: contig of 1468 bp in length  
14029 14128: gap of unknown length  
14129 14129: contig of 1356 bp in length  
14129 15485: gap of unknown length  
15485 17253: contig of 1669 bp in length  
17254 17353: gap of unknown length  
17354 18799: contig of 1446 bp in length  
18800 18899: gap of unknown length  
18900 20239: contig of 1340 bp in length  
20240 20339: gap of unknown length  
20340 21494: contig of 1155 bp in length  
21495 21594: gap of unknown length  
21595 22660: contig of 1066 bp in length  
22661 22760: gap of unknown length  
22761 24315: contig of 1555 bp in length  
24316 24415: gap of unknown length  
24416 26021: contig of 1606 bp in length  
26022 26121: gap of unknown length  
26122 27257: contig of 1136 bp in length  
27258 27357: gap of unknown length

```
* 27358 28620: contig of 1263 bp in length
* 28621 gap of unknown length
* 28720: contig of 1233 bp in length
* 28721 gap of unknown length
* 29553: contig of 1233 bp in length
* 30053: gap of unknown length
* 32125: contig of 2072 bp in length
* 32126 gap of unknown length
* 32225: contig of 1472 bp in length
* 32226 gap of unknown length
* 33698: contig of 1562 bp in length
* 33699 gap of unknown length
* 33798 35359: contig of 1562 bp in length
* 33799 gap of unknown length
* 35459: gap of unknown length
* 35460 37021: contig of 1562 bp in length
* 35461 gap of unknown length
* 37022 37121: gap of unknown length
* 37122 38194: contig of 1073 bp in length
* 38195 38294: gap of unknown length
* 38295 40095: contig of 1801 bp in length
* 40096 40195: gap of unknown length
* 40196 42087: contig of 1892 bp in length
* 42088 42187: gap of unknown length
* 42188 43765: contig of 1578 bp in length
* 43766 43866 gap of unknown length
* 43866 44931: contig of 1066 bp in length
* 44932 45031: gap of unknown length
* 45032 46278: contig of 1247 bp in length
* 46279 46379: gap of unknown length
* 46379 48544: contig of 2166 bp in length
* 48545 48644: gap of unknown length
* 48645 49741: contig of 1097 bp in length
* 49742 49841: gap of unknown length
* 49842 51329: contig of 1388 bp in length
* 51330 51329: gap of unknown length
* 51330 53356: contig of 2027 bp in length
* 53357 53456: gap of unknown length
* 53457 55035: contig of 1579 bp in length
* 55036 55135: gap of unknown length
* 55136 56770: contig of 1535 bp in length
* 56771 56770: gap of unknown length
* 56771 58151: contig of 1381 bp in length
* 58152 58251: gap of unknown length
* 58252 60521: contig of 2270 bp in length
* 60522 60622: gap of unknown length
* 60622 62884: contig of 2163 bp in length
* 62885 65094: contig of 2210 bp in length
* 65095 65194: gap of unknown length
* 65195 67815: contig of 2621 bp in length
* 67816 67915: gap of unknown length
* 67916 69794: contig of 1879 bp in length
* 69795 69894: gap of unknown length
* 69895 72567: contig of 2673 bp in length
* 72568 72667: gap of unknown length
* 72668 74508: contig of 1841 bp in length
* 74509 74608: gap of unknown length
* 74609 77519: contig of 2911 bp in length
* 77520 77619: gap of unknown length
* 77620 80278: contig of 2659 bp in length
* 80279 80378: gap of unknown length
* 80379 83008: contig of 2630 bp in length
* 83009 83108: gap of unknown length
* 83109 87336: contig of 4228 bp in length

Query Match      84.0%; Score 16.8; DB 2; Length 126056;
Best Local Similarity 90.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
Db 35984 CATGTCTGTGCTAGTCC 35965
|||||
|
```

```
RESULT 44
AC148290 132181 bp DNA linear PLN 07-MAY-2004
LOCUS Medicago truncatula chromosome 2 BAC clone mth2-45h12, complete
DEFINITION
```

```
ACCESSION AC148290
VERSION AC148290.3 GI:47077943
KEYWORDS HTG.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 132181)
Town, C.D., Tallon, L.J., Arbogast, T., Althoff, R., Hine, E.,
Monaghan, E., Smith, S.A., Utterback, T., Feldblyum, T., Koo, H. and
Cheung, F.
Medicago truncatula BAC genomic sequence
Unpublished
2 (bases 1 to 132181)
Town, C.D.
Direct Submission
Submitted (19-FEB-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 132181)
Town, C.D.
Direct Submission
Submitted (05-MAR-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
4 (bases 1 to 132181)
Town, C.D.
Direct Submission
Submitted (07-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
COMMENT On May 7, 2004 this sequence version replaced gi:45120164.
mononucleotide (T) repeat of 21 to 22 nucleotides.
FEATURES
source
1..132181
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/chromosome="2"
/clone="mth2-45h12"

ORIGIN
Query Match      84.0%; Score 16.8; DB 8; Length 132181;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
Db 47236 CGTGTCTGTGCTAGTCC 47255
|||||
|
```

```
RESULT 45
AC151300 134430 bp DNA linear HTG 25-AUG-2004
LOCUS Xenopus tropicalis clone CH216-107C14, WORKING DRAFT SEQUENCE, 2
DEFINITION Xenopus tropicalis
unpublished pieces.
ACCESSION AC151300
VERSION AC151300.1 GI:51536642
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 134430)
DOE Joint Genome Institute.
Unpublished
2 (bases 1 to 134430)
DOE Joint Genome Institute.
Direct Submission
Submitted (25-AUG-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
```

```

COMMENT
94598-1698, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 3608807
Center clone name: CHORI-216_107C14
-----
Summary Statistics
Consensus quality: 133976 bases at least Q40
Consensus quality: 134224 bases at least Q30
Consensus quality: 134293 bases at least Q20
Estimated insert size: 133000; agarose-fp estimation
Estimated insert size: 134330; sum-of-contigs estimation
Quality coverage: 11.15 in Q20 bases; agarose-fp estimation
Quality coverage: 11.04 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 50955: contig of 50955 bp in length
* 50956 51055: gap of unknown length
* 51056 134430: contig of 83375 bp in length.
Location/Qualifiers
1. .134430
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="CH216-107C14"
/clone_lib="CHORI-216 Xenopus tropicalis BAC library"

ORIGIN
Query Match      84.0%; Score 16.8; DB 2; Length 134430;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGTGTCTGCTGCTAGTCCC 20
|||||
Db 62193 CATGTGTCCTGCTAGTCCC 62212

RESULT 46
AC145410/c
LOCUS          147434 bp      DNA      linear      HTG 06-SEP-2003
DEFINITION    Bos taurus clone RP42-227G24, WORKING DRAFT SEQUENCE, 3 ordered
              pieces.
ACCESSION     AC145410
VERSION       AC145410.2 GI:34494894
KEYWORDS      HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE        Bos taurus (cow)
ORGANISM      Bos taurus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
              Bovinae; Bos.
              1 (bases 1 to 147434)
              Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
              Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
              Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
              Gupta,J., Haghigbi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
              Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-O.,
              Legaspi,R., McDuro,Q.L., Maduro,V.B., Margulies,E.H., Masfello,C.,
              Maskeri,B., McDowell,J., Paquirigan,C., Pearson,R., Portnoy,M.E.,
              Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
              Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
              Vogt,J.L., Wetherby,K.D., Young,A. and Green,E.D.
              NISC Comparative Sequencing Initiative
              Unpublished

REFERENCE
AUTHORS
2 (bases 1 to 147434)
Green,E.D.
Direct Submission
Submitted (02-JUL-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 147434)
Green,E.D.
Direct Submission
Submitted (06-SEP-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Sep 6, 2003 this sequence version replaced gi:32401574.
-----
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
-----
Project Information
Center project name: eyb
Center clone name: 227G24

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 147115 bases at least Q40
Consensus quality: 147204 bases at least Q30
Consensus quality: 147222 bases at least Q20
Insert size: 145000; agarose-fp
Insert size: 147234; sum-of-contigs
Quality coverage: 12.5ix in Q20 bases; agarose-fp
Quality coverage: 12.32x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11857: contig of 11857 bp in length
* 11858 11957: gap of unknown length
* 11958 140052: contig of 128095 bp in length
* 140053 140152: gap of unknown length
* 140153 147434: contig of 7282 bp in length.
Location/Qualifiers
1. .147434
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-227G24"
/clone_lib="RP42"
1. .11857
/clone="assembly_fragment"
/clone_end:SP6
vector_side:left
11958..140052
/clone="assembly_fragment"
133983..147434
/clone="clone overlaps with GenBank Accession Number

FEATURES
source
misc_feature
misc_feature
misc_feature

```



Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Macdonald,P., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

# TITLE

JOURNAL Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 5 (bases 1 to 159893)

## AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

# TITLE

JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Jul 30, 2002 this sequence version replaced gi:21426344.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L26195

Center clone name: L26195

-----

## FEATURES

### source

Location/Qualifiers  
1. .159893  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="15"  
/map="15"  
/clone="RP11-152L20"  
/clone\_lib="RPCT-11 Human Male BAC"  
423..436  
/note="single clone coverage"  
complement(2228..3572)  
/rpt\_family="L2"  
3603..3632  
/rpt\_family="AT\_rich"  
complement(3943..4111)  
/rpt\_family="LIP4"  
4112..4161

### unsure

repeat\_region  
complement(2228..3572)  
repeat\_region  
3603..3632  
repeat\_region  
complement(3943..4111)  
repeat\_region  
4112..4161

Query Match 84.0%; Score 16.8; DB 9; Length 159893;  
Best Local Similarity 90.0%; Pred. No. 4.4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTCTAGTCCC 20  
|||||  
Db 23875 CGTGTCTGTCTCAAGTCTC 23894

## RESULT 49

AC116162

LOCUS

AC116162

160580 bp

DNA

linear

PRI 29-DEC-2002

repeat\_region  
4506..4541  
/rpt\_family="LIP4"  
complement(5160..5217)  
repeat\_region  
5218..5362  
/rpt\_family="L2"  
repeat\_region  
5363..5643  
/rpt\_family="TA)n"  
complement(5654..5766)  
repeat\_region  
5813..5965  
/rpt\_family="MLT1J"  
complement(5972..6167)  
repeat\_region  
6168..6736  
/rpt\_family="LIME2"  
complement(6745..6780)  
repeat\_region  
6781..6815  
/rpt\_family="MLT2D"  
6816..6842  
/rpt\_family="(TG)n"  
6843..7183  
/rpt\_family="(GA)n"  
complement(7186..7210)  
repeat\_region  
7211..7499  
/rpt\_family="LIME2"  
complement(7500..7794)  
repeat\_region  
7795..8360  
/rpt\_family="LIME2"  
8361..8399  
/rpt\_family="AT\_rich"  
complement(8492..8638)  
repeat\_region  
8639..9875  
/rpt\_family="FLAM\_C"  
complement(9876..9978)  
repeat\_region  
9979..10101  
/rpt\_family="LIME2"  
10102..11071  
/rpt\_family="HAL1"  
11072..11305  
/rpt\_family="HAL1"  
complement(11422..11607)  
repeat\_region  
11608..12455  
/rpt\_family="ERVL"  
complement(12456..12781)  
repeat\_region  
12782..13174  
/rpt\_family="ERVL"  
complement(13175..13455)  
repeat\_region  
13456..13884  
/rpt\_family="ERVL"  
complement(13885..14489)  
repeat\_region  
14490..15187  
/rpt\_family="LIME"  
complement(15188..15624)



Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L26182  
 Center clone name: 268\_L\_17  
 -----

## FEATURES

```

source
1..160580
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="15"
  /map="15"
  /clone="RP11-268L17"
  /clone_lib="RPC1-11 Human Male BAC"
  1558..1589
    /rpt_family="CGG)n"
  1627..1680
    /rpt_family="C-rich"
  2078..2110
    /rpt_family="G-rich"
  complement(3581..3697)
    /rpt_family="L2"
  complement(5081..5216)
    /rpt_family="MIR"
  6033..6061
    /rpt_family="(TTTG)n"
  complement(6422..6548)
    /rpt_family="L2"
  complement(6549..6652)
    /rpt_family="MIR"
  6791..7072
    /rpt_family="MER63"
  complement(7460..7929)
    /rpt_family="MLT1J"
  8499..8703
    /rpt_family="MIR"
  8750..8773
    /rpt_family="AT-rich"
  8774..8887
    /rpt_family="LIME4A"
  9095..9145
    /rpt_family="L2"
  13294..14510
    /rpt_family="HSMAR2"
  14515..14696
    /rpt_family="(TTC)n"
  14697..14765
    /rpt_family="(TCC)n"
  complement(14768..14900)
    /rpt_family="FLAM_A"
  14905..14954
    /rpt_family="HSMAR2"
  15070..15158

```

Query Match 84.0%; Score 16.8; DB 9; Length 160580;  
 Best Local Similarity 90.0%; Pred. No. 4.4e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 CGTGTGCTGCTAGTCCC 20
   |||||
DB 128514 CGTGTGCTGCTAGTCCC 128533

```

## RESULT 50

```

AP006203
LOCUS      168921 bp DNA linear PRI 18-FEB-2003
DEFINITION Homo sapiens genomic DNA, chromosome 8, clone:RP11-110N16, complete
            sequence.
ACCESSION AP006203
VERSION   AP006203.1 GI:28411639
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,  
 Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Homo sapiens genomic DNA  
 Published Only in Database (2003)  
 2 (bases 1 to 168921)  
 Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,  
 Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Direct Submission  
 Submitted (14-FEB-2003) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/  
 Tel.81-45-503-9111, Fax:81-45-503-9170)  
 FEATURES  
 Location/Qualifiers  
 1..168921  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /clone="RP11-110N16"

## ORIGIN

```

Query Match 84.0%; Score 16.8; DB 9; Length 168921;
Best Local Similarity 90.0%; Pred. No. 4.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGCTAGTCCC 20
   |||||
DB 73083 CGTGTGCTGCTAGTCCC 73102

```

Search completed: March 15, 2005, 19:01:07  
 Job time : 1721 secs

**THIS PAGE BLANK (USPTO)**